



Universidade de Brasília

Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas

Programa de Pós-Graduação em Administração

DIOGO RIBEIRO DA FONSECA

**Avaliação da Eficiência de Unidades de Treinamento da
Administração Pública Federal**

Brasília – DF

2021

DIOGO RIBEIRO DA FONSECA

**Avaliação da Eficiência de Unidades de Treinamento da
Administração Pública Federal**

Tese submetida ao Programa de Pós-Graduação em Administração da
Universidade de Brasília como requisito para a obtenção do título de
Doutor em Administração.

Banca Examinadora:

Prof. Dr. Pedro Paulo Murce Menezes Cavalcante
Orientador (Universidade de Brasília – PPGA/UnB).

Prof^a. Dr^a. Gardênia da Silva Abbad
Examinadora Interna (Universidade de Brasília – PPGA/UnB).

Prof. Dr. José Dionisio Gomes da Silva
Examinador Externo (Universidade Federal do Rio Grande do Norte – UFRN).

Prof^a. Dr^a. Eliane Maria Pires Giavina Bianchi
Examinadora Externa (Centro Universitário Campo Limpo Paulista – UNIFACCAMP).

Prof. Dr. Carlos Rosano Peña
Examinador Suplente (Universidade de Brasília – PPGA/UnB).

Brasília, 30 de agosto de 2021

Agradecimentos

A conclusão dessa Tese, em tempos tão difíceis, jamais seria possível sem o apoio, o amor e a dedicação de minha eterna e amada pessoa favorita para absolutamente todos os momentos de vida, Luciana Graziani, a quem agradeço profundamente por ter estado ao meu lado me inspirando e incentivando, diariamente, a superar novos limites e dificuldades e a acreditar que, afinal de contas, tudo é possível, quando deixamos de pensar que estamos sós. Obrigado, amor!

Agradeço ao meu pai, Aleilton, minha mãe, Rosana, e meu irmão, Raul, pelo porto seguro, pelo zelo e pelo amor incondicional que, mesmo de longe há tanto tempo, me trazem segurança e ímpeto em meio às trilhas da vida profissional e acadêmica, nem sempre fáceis de cruzar.

Agradeço a todos os meus queridos amigos de longa data, André, Elisa, Anna, Gil, Guto, Gabriel, Bawdy, Monique, que apesar de estarem todos, todos, vivendo longe e me causando imensas saudades, estiveram genuinamente na torcida todos esses anos, sendo infinitamente compreensivos e sempre me lembrando de nossos futuros encontros e tudo que ainda temos para ver e viver juntos, sem ninguém jamais “soltar a mão” de ninguém.

Ao meu precioso segundo lar, junto à Dona Regina, Dona Rita e à família Graziani, onde o acolhimento sempre existiu, e que eu também acolherei com todo meu coração pelo resto de minha vida.

Agradeço a todos os meus colegas e amigos do setor público que, assim como eu, compartilham de um profundo desejo de tornar real tudo aquilo que acreditamos ser possível para tornar melhor e mais forte o serviço público no Brasil. Sem vocês não haveria motivo para a minha vida acadêmica. Muitos de vocês são meus parceiros de pesquisa há quase uma década e sempre serei grato por todas as vezes que pude contar com o apoio de cada um nessa trajetória.

Essa trajetória não teria sido possível sem os professores de antes e de agora, aos quais agradeço imensamente todos os ensinamentos, lições e experiências que me trouxeram até aqui. Assim, agradeço ao Prof. Pedro Menezes pela confiança e pelo crescimento que me proporcionou em todos esses anos de pesquisa. Agradeço também ao Prof. Rosano pelo generoso e sincero apoio nos meandros dessa pesquisa, me permitindo compreender os recursos que eu tinha ao meu alcance.

Agradeço à Professora Gardênia Abbad pela inspiração de uma nova perspectiva de análise que deu um novo sentido a esse trabalho muito além do que eu poderia enxergar.

Agradeço imensamente à Professora Eliane Bianchi e ao Professor José Dionísio que, muito gentilmente, aceitaram o convite para compor a banca para avaliação desse trabalho e contribuíram com valiosas perspectivas para esse trabalho.

Por fim, tive o privilégio de atuar como professor nesses anos de doutorado e sou imensamente grato a cada um dos meus alunos pelo crescimento que essa belíssima prática me proporcionou e que me brindou com a honra de acompanhá-los em suas trajetórias para que se tornassem, hoje, colegas de profissão. Sempre contem comigo.

Resumo

O presente estudo teve como objetivo principal propor um método de avaliação da eficiência de unidades de treinamento na Administração Pública Federal brasileira, tendo por consideração critérios de economicidade e alcance dos programas de treinamento realizados em organizações públicas federais. A relevância da pesquisa justifica-se pela necessidade de complementação das perspectivas orientadas para a avaliação dos resultados de aprendizagem, consolidadas no campo de TD&E, ao permitir um olhar sob uma perspectiva gerencial dessas atividades, abrangendo a otimização das decisões de alocação dos recursos de treinamento. Para o alcance do objetivo proposto, buscou-se operacionalizar um modelo de Análise Envoltória de Dados (DEA) para cálculo de índices de eficiência relativa das unidades de treinamento e análise da influência de fatores técnicos, setoriais e organizacionais nos seus resultados. A partir de uma revisão sistemática de literatura foram definidas e testadas variáveis selecionadas para a constituição de um modelo Network-DEA em duas etapas. Os dados foram coletados junto a seis organizações públicas federais, compreendendo informações sobre processos e ações de treinamento realizados no período de 2015 a 2019. A análise foi realizada por meio de abordagem intertemporal, comparando as seis organizações ao longo dos cinco anos, totalizando uma amostra de 30 casos. Por fim, variáveis discricionárias e não-discricionárias referentes à atividade de treinamento e à estrutura organizacional foram analisadas para a estimativa de suas influências na eficiência das unidades de treinamento. Como principal contribuição, o método proposto permite o cálculo de índices de eficiência global para a avaliação e classificação de unidades de treinamento com relação à sua capacidade de provisão de cursos para o maior número de servidores ao menor custo. Nesse sentido, o estudo alcançou seu objetivo de contribuir com a literatura de métodos de avaliação de treinamento ao propor uma perspectiva voltada para a eficiência com foco no processo gerencial de alocação dos recursos. O presente trabalho estabelece uma integração entre os campos da Pesquisa Operacional, Economia, TD&E e Gestão Pública de forma a promover contribuições a cada um deles pela utilização de perspectivas teóricas, metodológicas e de escopo de análise que são, em geral, pouco exploradas. Consideram-se limitações do estudo a dimensão da amostra e o enfoque operacional do modelo de análise, sendo recomendada a realização de estudos futuros em amostras mais extensas que viabilizem o aperfeiçoamento das técnicas de análise, assim como a ampliação do modelo com a inclusão de outras variáveis ambientais e sobretudo de variáveis de resultados de treinamento para indivíduos, organizações e sociedade.

Palavras-chave: Treinamento; Avaliação; Eficiência; Análise Envoltória de Dados; Setor Público.

Abstract

The main objective of this study is to provide a method for evaluating the efficiency of training programs in the Brazilian Federal Public Administration, using as criteria the lowest possible costs and the greater equity of training provision in public organizations. The research is justified by its contribution as a distinct managerial perspective, regarding resource allocation decisions in training activities, in relation to the long-established literature of training evaluation which emphasizes the training effectiveness perspective for the achievement of individual and organizational learning outcomes. Following this research objective, the Data Envelopment Analysis (DEA) method was conceived as a means for the calculation of relative efficiency scores of training programs and for the analysis of technical, managerial and organizational influences on training efficiency results. A literature review found several variables for the definition and testing of various DEA models, resulting in the proposition of a two-stage Newotk-DEA model for training efficiency. After a preliminary account of the research field, data was collected from six federal public agencies in order to evaluate comparatively the efficiency of their training programs. Data collected comprised information regarding which processes and training activities were carried out in the organizations during the years of 2015 to 2019. The DEA analysis was carried out with an intertemporal approach, comparing the six organizations over this five period year, totaling a sample of 30 cases. Based on calculated efficiency scores, a two-step analysis sought to evaluate the effects of conditioning factor such as the activities of training processes and organizational features on the efficiency of training units. The main contribution of this study refers to the calculation of a global efficiency score, which allows for an adequate evaluation and classification of training units in relation to their capacity to deliver courses to a greater number of civil servants at the lowest costs. Hence, the study achieved its objectives of contributing to the literature on training evaluation with a method for providing an efficiency approach to study optimal resources combination and allocation in training. In addition, this work also establishes an integration between the fields of Operational Research, Economy, T&D and Public Management in order to promote contributions to each field through a relatively unexplored analytical scope. The limitations of this study regard the relatively small sample of organizations as well as the more restrict number of variables. Future studies should consider a larger sample of organizations that enable a larger set of analytical techniques, as well incorporate new variables to the model allowing for the evaluation of training efficiency considering more environmental factors and, mainly, variables regarding results criteria on individual, organizational and societal levels.

Keywords: Training; Evaluation; Efficiency; Data Envelopment Analysis; Public sector.

Lista de Figuras

Figura 1. Relação entre conceitos e respectivas soluções educacionais.....	28
Figura 2. Sistema de Treinamento.....	31
Figura 3. Ilustração da cadeia de objetivos em um programa de educação em saúde.....	38
Figura 4. Níveis de Avaliação em TD&E	45
Figura 5. Modelo de Avaliação e Pesquisa em Desenvolvimento de Recursos Humanos.....	55
Figura 6. Uma abordagem insumo-processo-produto para a avaliação de treinamento.....	59
Figura 7. Modelo de avaliação integrado e somativo - MAIS.....	61
Figura 8. Modelo IMPACT de Avaliação do Treinamento.....	63
Figura 9. O Modelo Lógico básico.....	66
Figura 10. Estrutura do Modelo Lógico	67
Figura 11. Exemplo de Modelo Lógico Final	69
Figura 12. Mapa Estratégico do Treinamento	73
Figura 13. Integração de Modelos Sistêmicos para avaliação da eficiência em TD&E.....	89
Figura 14. Visão sistêmica para avaliação da eficiência	91
Figura 15. Função Produção	97
Figura 16. Tipos de eficiência	101
Figura 17. Representação de uma DMU	104
Figura 18. Métodos de Análise de Eficiência.....	105
Figura 19. Representação de uma DMU em série de três estágios.....	108
Figura 20. Orientação dos modelos matemáticos para os inputs e os outputs.....	111
Figura 21. Comparação entre as fronteiras dos modelos BCC e CCR.....	114
Figura 22. Modelo DEA para avaliação de programas de pós-graduação	131
Figura 23. Modelo DEA para avaliação de universidades de pesquisa.....	132
Figura 24. Sumário de procedimentos de busca e filtragem de artigos.....	138
Figura 25. Modelo Conceitual de Análise em dois estágios.....	159
Figura 26. Modelo inicial	161
Figura 27. Quantidade de participações x percentual da força de trabalho capacitada.....	166
Figura 28. Modelo Final de Análise	169
Figura 29. Modelo Aditivo	173
Figura 30. Evolução dos índices de eficiência por período.....	178
Figura 31. Fronteira de Eficiência CCR x BCC para o Estágio 1	182

Figura 32. Fronteira de Eficiência CCR x BCC corrigida para o Estágio 1	184
Figura 33. Eficiência e Intervalos de confiança para o Estágio 1	185
Figura 34. Fronteira de Eficiência BCC sem orientação específica para o Estágio 2	191
Figura 35. Fronteira de Eficiência BCC corrigida para o Estágio 2	192
Figura 36. Eficiência e Intervalos de confiança para o Estágio 2	193
Figura 37. Gráfico Eficiência global x Força de Trabalho capacitada	200
Figura 38. Gráfico Eficiência global x Custo da hora-aula	200
Figura 39. Gráfico Equipe x Eficiência global	202
Figura 40. Gráfico Recursos financeiros x Eficiência global	202
Figura 41. Eficiência e proporção de gastos por modalidade	204
Figura 42. Gráfico Eficiência global x Treinamentos presenciais externos	205
Figura 43. Gráfico Eficiência Global x Treinamento EaD	206
Figura 44. Gráfico Eficiência global x Treinamentos presenciais internos	206
Figura 45. Eficiência e Variáveis contextuais da Organização A	215
Figura 46. Eficiência e Variáveis contextuais da Organização B	218
Figura 47. Eficiência e Variáveis contextuais da Organização C	220
Figura 48. Eficiência e Variáveis contextuais da Organização D	223
Figura 49. Eficiência e Variáveis contextuais da Organização E	225
Figura 50. Eficiência e Variáveis contextuais da Organização F	227

Lista de Tabelas

Tabela 1. Principais modelos de avaliação de treinamento do campo de TD&E.....	48
Tabela 2. Critérios de modelos de avaliação em TD&E em abordagem sistêmica.....	80
Tabela 3. Tipos de Eficiência	102
Tabela 4. Variáveis de insumos utilizadas para avaliação da eficiência educacional	124
Tabela 5. Variáveis de produtos utilizadas para avaliação da eficiência educacional	126
Tabela 6. Variáveis não-discrecionárias utilizadas para avaliação da eficiência educacional	128
Tabela 7. Relação de insumos e produtos utilizados em eficiência na Educação	129
Tabela 8. Termos de busca exploratória em bases de dados	136
Tabela 9. Variáveis encontradas na revisão de literatura	140
Tabela 10. DMUs de artigos com abordagem DEA.....	143
Tabela 11. Principais produtos (outputs) utilizados nos modelos DEA da amostra.....	145
Tabela 12. Principais insumos utilizados nos modelos DEA da amostra.....	147
Tabela 13. Organizações selecionadas	158
Tabela 14. Teste de Correlação de variáveis	167
Tabela 15. Evolução da Eficiência Técnica e de Escala para o Estágio 1.....	183
Tabela 16. Análise de Benchmarks para o Estágio 1	187
Tabela 17. Análise de Supereficiência para o Estágio 1	189
Tabela 18. Evolução da Eficiência Técnica e de Escala para o Estágio 2.....	191
Tabela 19. Análise de Benchmarks para o Estágio 2	194
Tabela 20. Análise de Supereficiência para o Estágio 2	195
Tabela 21. Eficiência global e combinações de insumos e produtos.....	196
Tabela 22. Eficiência global e indicadores do processo de treinamento	198
Tabela 23. Efeitos das variáveis não-discrecionárias nas eficiências	210

Lista de Abreviaturas e Siglas

- ANT** – Avaliação de Necessidades de Treinamento
- ASTD** – *American Society for Training and Development*
- BSC** – *Balanced Scorecard*
- CHA** – Conhecimentos, habilidades e atitudes
- CRS** – *Constant Returns to Scale*
- DASP** – Departamento de Administração do Serviço Público
- DEA** – Análise Envoltória de Dados
- DMU** - *Decision Making Unit*
- DRH** – Desenvolvimento de Recursos Humanos
- FDH** – Análise de Livre Disponibilidade
- FI** – Fator de impacto
- GPC** – Gestão por Competências
- IES** – Instituição de Ensino Superior
- KS** – Teste de Kolmogorov-Smirnov
- MBA** – *Master in Business Administration*
- P&D** – Pesquisa e Desenvolvimento
- PNDP** – Política Nacional de Desenvolvimento de Pessoal
- ROI** – Retorno sobre investimento
- SFA** – Análise de Fronteiras Estocásticas
- SGP/ME** – Secretaria de Gestão e Desempenho de Pessoal do Ministério da Economia
- SIPEC** – Sistema de Pessoal Civil da Administração Federal
- SJR** – SCImago Journal Rank
- TD&E** – Treinamento, Desenvolvimento & Educação
- TEC** – *Training and Enterprise Council*
- T&D** – Treinamento e Desenvolvimento
- TGS** – Teoria Geral dos Sistemas
- VRS** – *Variable Returns to Scale*

SUMÁRIO

1	Introdução	11
1.1	Objetivo geral.....	16
1.2	Objetivos específicos	16
1.3	Justificativa	16
2	Referencial Teórico.....	19
2.1	Aprendizagem e Treinamento no Contexto Organizacional	19
<i>2.1.1</i>	<i>Evolução do campo de TD&E.....</i>	<i>23</i>
<i>2.1.2</i>	<i>O conceito de Treinamento em TD&E.....</i>	<i>26</i>
<i>2.1.3</i>	<i>O Sistema de TD&E.....</i>	<i>31</i>
2.2	Fundamentos e Perspectivas da Avaliação de Programas.....	33
<i>2.2.1</i>	<i>Modelos de avaliação em TD&E.....</i>	<i>43</i>
<i>2.2.2</i>	<i>Perspectivas de avaliação em TD&E</i>	<i>77</i>
<i>2.2.3</i>	<i>A perspectiva da eficiência na avaliação em TD&E</i>	<i>87</i>
2.3	Avaliação da eficiência	94
<i>2.3.1</i>	<i>O conceito de eficiência.....</i>	<i>96</i>
<i>2.3.2</i>	<i>Tipos de eficiência</i>	<i>100</i>
<i>2.3.3</i>	<i>Métodos de Avaliação da Eficiência</i>	<i>103</i>
<i>2.3.4</i>	<i>Análise Envoltória de Dados</i>	<i>109</i>
<i>2.3.5</i>	<i>Modelagem e Seleção de Variáveis em DEA</i>	<i>115</i>
2.4	Avaliação da eficiência em Educação	118
2.5	Revisão de literatura: Eficiência das atividades de treinamento	133
<i>2.5.1</i>	<i>Análise das variáveis utilizadas nos artigos empíricos selecionados.....</i>	<i>139</i>
<i>2.5.2</i>	<i>Artigos com aplicação do DEA</i>	<i>143</i>
3	Método	151
3.1	Universo de Pesquisa	151
3.2	Caracterização da pesquisa	152
3.3	Pesquisa preliminar de campo.....	153
3.4	Instrumentos de pesquisa.....	155
3.5	Procedimentos e técnicas de coleta de dados.....	156
3.6	Amostra de organizações pesquisadas	157
3.7	Definição do modelo e variáveis de análise.....	158
<i>3.7.1</i>	<i>Extensões ao modelo DEA adotadas.....</i>	<i>159</i>
<i>3.7.2</i>	<i>Teste de modelos e variáveis selecionadas</i>	<i>161</i>

3.7.3	<i>Modelo e Base de Dados Final</i>	168
3.8	Procedimentos e Técnicas de Análise	169
3.8.1	<i>Análise Intertemporal de dados em Painel</i>	169
3.8.2	<i>Análise da eficiência e definição de pressupostos de análise</i>	170
3.8.3	<i>Outliers e Supereficiência</i>	171
3.8.4	<i>DEA Bootstrap</i>	172
3.8.5	<i>Modelos aditivos, análise de folgas (slacks) e benchmarks</i>	172
3.8.6	<i>Análise de variáveis não-discrecionárias: Modelo de dois estágios</i>	174
3.8.7	<i>Interpretação dos índices de eficiência</i>	174
4	Resultados e discussão	176
4.1	Avaliação conjunta da eficiência das unidades de treinamento	177
4.2	Avaliação comparativa da eficiência das unidades de treinamento	181
4.2.1	<i>Estágio 1: Produção</i>	181
4.2.2	<i>Estágio 2: Entrega</i>	189
4.2.3	<i>Eficiência Global</i>	196
4.3	Análise de fatores interferentes na eficiência das unidades de treinamento	208
4.4	Recomendações e implicações para as organizações	214
4.4.1	<i>Organização A</i>	215
4.4.2	<i>Organização B</i>	218
4.4.3	<i>Organização C</i>	220
4.4.4	<i>Organização D</i>	223
4.4.5	<i>Organização E</i>	225
4.4.6	<i>Organização F</i>	227
5	Considerações finais	230
	Referências	235
	APÊNDICE A - Lista de artigos da Revisão de Literatura	251
	APÊNDICE B – Carta-convite	255
	APÊNDICE C - Formulário de Descrição de Equipe	258
	APÊNDICE D – Formulário de Descrição de Processos de Capacitação	259
	APÊNDICE E – Base de dados final	262
	APÊNDICE F – Base de dados: Fatores não discrecionários dicotomizados	264
	APÊNDICE G - Script final no software R	266
	APÊNDICE H - Script final no software R: variáveis não-discrecionárias	275

1 Introdução

Diante de ambientes cada vez mais dinâmicos e complexos, a qualificação contínua da força de trabalho tornou-se fundamental no ambiente organizacional. Nesse contexto, as atividades de Treinamento, Desenvolvimento e Educação (TD&E) adquiriram grande relevância como uma função estratégica nas organizações contemporâneas. Conseqüentemente, as atividades de treinamento se expandiram nas organizações por meio da criação de departamentos especializados voltados para o aprimoramento do desempenho individual dos trabalhadores, mas também para a melhoria do desempenho organizacional e para geração de benefícios à sociedade, na medida em que apoia o desenvolvimento de vantagens competitivas e inovações que promovem maior crescimento econômico e desenvolvimento social (Aguinis & Kraiger, 2009; Ferreira & Abbad, 2014; Kraiger, McLinden & Casper, 2004; Meneses & Abbad, 2009; Pilati, 2006; Salas, Tannenbaum, Kraiger & Smith-Jentsch, 2012; Salas & Cannon-Bowers, 2001).

No contexto do setor público brasileiro em específico, políticas para profissionalização e qualificação de servidores públicos foram compreendidas como fundamentais para a profissionalização da burocracia, tendo em vista a adequada provisão dos serviços públicos à sociedade. As primeiras políticas de desenvolvimento de servidores públicos podem ser traçadas desde a criação do Departamento de Administração do Serviço Público - DASP em 1938 (Cavalcante & Carvalho, 2017; Pacheco, 2002). Não obstante, ao longo desse período de consolidação da burocracia pública brasileira, priorizou-se o aperfeiçoamento de outros subsistemas de gestão de pessoas, de forma que as políticas de formação e capacitação de servidores públicos se desenvolveram de maneira assistemática, com foco na progressão na carreira, levando à não constituição de uma política abrangente para o desenvolvimento permanente de servidores públicos (Pacheco, 2002).

No ano de 2006, como um meio de organizar, aprimorar e ampliar os resultados dos processos de capacitação realizados nos órgãos, tornando mais efetiva a oferta de cursos realizados por escolas de governo (Camões & Meneses, 2016), o governo lançou o Decreto nº 5.707/2006, que instaurou a Política Nacional de Desenvolvimento de Pessoal – PNDP. O Decreto definiu como objetivos expressos a melhoria dos serviços públicos por meio do desenvolvimento permanente de servidores, o alinhamento das competências individuais aos resultados de governo, a ampliação da divulgação de ações de capacitação e a racionalização e efetividade dos gastos com treinamento. Além disso, especificou a gestão por

competências como ferramenta gerencial, adotando uma terminologia de orientação estratégica e denotando um marco para a mudança paradigmática de uma gestão de pessoas funcional para um modelo estratégico de desenvolvimento de servidores públicos (Brasil, 2006).

A partir da edição do referido Decreto nº 5.707/2006, variadas iniciativas de implementação da gestão por competências e das estratégias de gestão da capacitação enunciadas pelo normativo ocorreram em diferentes organizações públicas. Ainda assim, com o decorrer de 13 anos após a sua promulgação, estudos demonstraram que a PNDP não foi implementada com sucesso nas organizações (Camões & Meneses, 2016; TCU, 2016). Diante desse cenário, em 2019, foram reformuladas as diretrizes da PNDP por meio da edição da mais recente política para desenvolvimento de pessoal na administração pública federal (ENAP, 2020).

A nova Política Nacional de Desenvolvimento de Pessoas, instituída com a promulgação do decreto nº. 9.991/2019, resulta da avaliação crítica da PNDP por organismos nacionais e internacionais (e.g. OCDE, 2010) e a apropriação de modelos em outros países incorporadas pelo governo brasileiro (ENAP, 2020). Dentre seus principais objetivos, o Decreto n.º 9.991/2019 estabelece o alinhamento das necessidades de desenvolvimento com a estratégia do órgão; o estabelecimento de objetivos e metas institucionais como referenciais para o planejamento da capacitação; a provisão de treinamento de maneira equânime entre servidores; a análise do custo-benefício do treinamento visando o adequado uso dos recursos públicos de acordo com os princípios da economicidade e da eficiência; e o monitoramento e avaliação das ações de desenvolvimento a fim de promover a racionalidade dos gastos públicos (Brasil, 2019a).

Os citados avanços na formulação e implementação da PNDP têm se proposto a superar as limitações e dificuldades observadas nos anos anteriores. Em complemento, estudos recentes vêm demonstrando que parte relevante das organizações públicas realizam ações educacionais de maneira assistemática e movidas por motivações e interesses dos indivíduos, nem sempre alinhadas às necessidades organizacionais (Camões & Meneses, 2016; Fonseca & Meneses, 2018; TCU, 2013; 2016), o que resulta no uso inadequado de recursos públicos com atividades que não contribuem efetivamente para os resultados de governo. Assim, a otimização do uso dos recursos públicos destinados à capacitação de servidores, que constitui princípio central não apenas da atual Política Nacional de Desenvolvimento de Pessoas, mas também do arcabouço normativo mais amplo do setor

público desde 1998, permanece como um objetivo não alcançado das unidades de treinamento. Dessa forma, a avaliação das ações de capacitação promovidas pelas organizações públicas constitui demanda urgente a fim de garantir melhores resultados e maior eficiência dos gastos.

Conforme explicitado por Meneses, Zerbini & Abbad (2010), a importância, complexidade e a dimensão que o TD&E possui nas organizações propiciou uma grande quantidade de pesquisas nessa área e a consequente consolidação de um campo científico específico denominado de Treinamento, Desenvolvimento e Educação (TD&E). Considerado como uma ciência em si, o campo de TD&E conta com um amplo arcabouço teórico, técnico e científico para a análise, planejamento, execução e avaliação de ações educacionais no contexto das organizações (Abbad, Pilati & Pantoja, 2003; Bell, Tannenbaum, Ford, Noe & Kraiger, 2017; Borges-Andrade, 2006a; 2006b; Pilati, 2006; Salas & Cannon-Bowers, 2001; Salas et al., 2012). Dentre os métodos desenvolvidos pelo campo para o aperfeiçoamento das atividades de treinamento nas organizações, as avaliações tornaram-se uma das principais ferramentas para determinação do valor e dos resultados das ações educacionais para os indivíduos e para a organização de forma a justificar os volumes de investimentos realizados nessa atividade (Abbad, Nascimento & Gaspar, 2020). Dessa forma, para a concretização dos objetivos de avaliação das ações educacionais no setor público, o campo de TD&E congrega as ferramentas metodológicas necessárias para a melhoria da eficiência e efetividade das políticas de desenvolvimento de servidores públicos.

A Avaliação em TD&E pode ser compreendida como a coleta e análise de dados com vistas ao julgamento de valor sobre a eficiência, eficácia e a efetividade de ações, programas ou projetos instrucionais (Abbad et al., 2020). Os primeiros modelos de avaliação de treinamento propostos por Kirkpatrick (1976) e Hamblin (1978) foram construídos em torno de cinco níveis tradicionais para a avaliação de resultados de treinamento, quais sejam: i) Reação, relativa à satisfação, percepção de utilidade e opiniões dos participantes acerca do treinamento logo após a sua realização; ii) Aprendizagem, que se refere ao nível de conhecimentos adquiridos; iii) Comportamentos, que diz respeito à transferência dos conhecimento para as atividades de trabalho; iv) Resultados, referente aos impactos para a organização; e v) Valor final, referente a resultados mais amplos para a organização e para a sociedade como um todo no longo prazo. Foi a partir dessa teoria que se estabeleceram os fundamentos para a maioria dos modelos de avaliação em treinamento que se seguiram.

Ao longo de seu desenvolvimento, fundamentando-se nas proposições de Gagné (1970), o campo de TD&E passou a adotar uma perspectiva sistêmica de forma que novos elementos, até então próprios de modelos de avaliação de outras disciplinas, tais como os modelos de contexto-insumo-processo-produto (Stufflebeam, 2003), foram incorporados aos modelos de avaliação de treinamento. Desta forma, as avaliações em TD&E assumiram um caráter integrativo e englobaram critérios relativos aos contextos em que estão inseridas, assim como variáveis relativas aos insumos, processos e características individuais dos programas e ações de treinamento que impactam nos resultados e efeitos para os indivíduos e para a organização (Abbad et al., 2020; Borges Andrade et. al., 2012; Holton, 1996; Meneses et al., 2010; Meneses & Abbad, 2009; Passmore & Velez, 2015; Tharenou, 2010).

A partir desses marcos, uma grande quantidade de modelos de avaliação em TD&E foi desenvolvida na literatura (Passmore & Velez, 2015), abrangendo diferentes estruturas e perspectivas de análise. Apesar disso, de forma geral, os níveis propostos por Kickpatrick e Hamblin ainda permanecem como objetivos de avaliação nos âmbitos científico e profissional do TD&E (Alliger, Tannenbaum, Bennet Jr. & Shotland, 1997; Kennedy, Chyung, Winiecki & Brinkerhoff, 2014; Passmore & Velez, 2015), de forma que estabeleceram a taxonomia fundamental de critérios em torno da qual a grande maioria dos modelos de avaliação em treinamento se estruturou (Abbad et al., 2020).

Nesse sentido, ainda que as avaliações tenham como finalidade a definição do mérito e valor das ações educacionais em critérios fundamentais como eficiência, eficácia e efetividade, o enfoque nos níveis tradicionais de resultados terminou por se tornar predominante dentre os modelos de avaliação em TD&E, de forma que a perspectiva da eficiência se tornou menos desenvolvida e abrangida pela maior parte dos modelos e práticas de avaliação. Ainda que, conforme Scriven (1991), se compreenda que o principal objetivo das avaliações deva de fato residir na eficácia e efetividade, a perspectiva da eficiência cumpre um papel complementar e distinto que se faz relevante para a análise completa de programas de treinamento, tendo um papel fundamental, por exemplo, na determinação das melhores alternativas para a utilização racional de recursos alocados a programas e atividades organizacionais.

A perspectiva da eficiência, portanto, está diretamente relacionada à solução de questões relativas à otimização dos gastos e investimentos em capacitação. As avaliações de eficiência, conforme Scriven (1991), para além dos resultados alcançados, se preocupam notadamente com a análise e otimização das decisões de alocação dos recursos, focando

sobretudo na análise do processo produtivo e não do impacto gerado (Rosano-Peña, 2008). Aplicado ao contexto do TD&E, o critério da eficiência visa garantir que os resultados das atividades de treinamento sejam os maiores possíveis com a menor utilização de recursos financeiros, humanos e materiais (Abbad et al., 2020).

Por esse motivo, assim como nos demais campos de conhecimento como delineado por Scriven (1991) e Suchman (1967), um programa integrado de avaliações deve buscar abranger conjuntamente os aspectos de eficiência, eficácia e efetividade. De forma geral, a eficácia e a efetividade de uma atividade, em termos dos critérios de valor, consistem nas principais dimensões (Scriven, 1991; Mark, Henry & Julnes, 2000). Por outro lado, uma vez constatados os resultados, a avaliação da eficiência permite a detecção de eventuais problemas operacionais e gerenciais, desperdícios de recursos e ineficiências alocativas que podem ter ocorrido e que, em última análise, podem impactar também os resultados que poderiam ter sido alcançados caso nenhum desses problemas estivesse presente. Nesse sentido, a avaliação da eficiência possibilita a complementação das perspectivas orientadas para a avaliação dos resultados de aprendizagem, ao permitir um olhar sob uma perspectiva gerencial para otimização de suas atividades.

Segundo Suchman (1967), as avaliações de eficiência possuem um importante papel na justificativa e dimensionamento de investimentos sobretudo no contexto do setor público em que, sob uma perspectiva normativa, a eficiência torna-se um princípio para os agentes públicos responsáveis pela utilização dos recursos para atendimento às necessidades coletivas (Modesto, 2000). Dessa forma, medidas de eficiência devem ser incorporadas sempre que possível na avaliação de programas e atividades organizacionais, visando justificar a continuidade dos investimentos realizados, em especial no setor público.

No entanto, atividades e programas organizacionais normalmente envolvem múltiplos insumos e produtos, além de aspectos contextuais que podem influenciar seus processos e resultados. Diante de questões dessa natureza, formularam-se no campo da Economia algumas definições específicas de eficiência, assim como técnicas e métodos para a avaliação de unidades produtivas complexas, com múltiplos insumos, produtos. Os modelos baseados no conceito de eficiência de Farrel (1957), por exemplo, buscaram o desenvolvimento de medidas para o cálculo da eficiência considerando-se diferentes fatores, tais como preços, escala e contexto. Dentre eles uma das abordagens mais utilizadas em diferentes campos, sobretudo para a avaliação da eficiência técnica e qualidade do gasto no setor público é o método da Análise Envoltória de Dados (*Data Envelopment Analysis* –

DEA) (Rosano-Peña, 2008; Sant'Anna, Lopes, Miranda, Bermejo & Demo, 2020; Siqueira, Farias, Bermejo & Sousa, 2020).

Assim, a questão que emerge reside na dificuldade da determinação e mensuração da eficiência de atividades complexas, de forma que se faz necessária a realização de estudos específicos a fim de compor modelos de avaliação capazes de abordar adequadamente tais aspectos.

1.1 Objetivo geral

Diante do exposto, esta pesquisa tem como objetivo principal propor um método de avaliação da eficiência de unidades de treinamento da Administração Pública Federal brasileira.

1.2 Objetivos específicos

Para a consecução do objetivo geral, foram realizados os seguintes objetivos específicos:

- Identificar variáveis que permitam a avaliação da eficiência de unidades de treinamento;
- Avaliar a eficiência de unidades de treinamento da Administração Pública Federal brasileira;
- Analisar o efeito de fatores interferentes na eficiência de unidades de treinamento da Administração Pública Federal brasileira.

1.3 Justificativa

Conforme Scriven (1991), uma avaliação consiste em um processo objetivo, sistemático e formal para determinação do valor de algo, tendo em vista critérios pré-definidos. No contexto das organizações, tais critérios podem se referir a diferentes tipos de resultados de programas e atividades organizacionais, envolvendo perspectivas de análise relacionadas à eficiência, eficácia e efetividade de suas atividades.

Relativamente à avaliação de treinamentos nas organizações, o TD&E é considerado um campo teoricamente consistente e consolidado no nível micro organizacional, abrangendo o desenvolvimento de tecnologia instrucional e a investigação de sua relação com variáveis individuais, tais como aprendizagem e desempenho, associadas à efetividade

das ações educacionais (Bell, Tannenbaum, Ford, Noe, & Kraiger, 2017; Salas, Tannenbaum, Kraiger, & Smith-Jentsch, 2012; Tharenou, 2010). No entanto, a realização de avaliações que tenham como foco a eficiência de ações e programas de treinamento ainda é relativamente pouco desenvolvida no campo (Passmore & Velez, 2015).

Nesse sentido, da mesma forma que o campo de TD&E estabelece com clareza quais as principais atividades e condições para a efetividade das ações de treinamento nas organizações (e.g. Salas et al., 2012), com relação à eficiência, no entanto, não há uma sistematização de quais atividades e condições promovem um melhor uso dos recursos de TD&E. Assim, um aprofundamento teórico acerca da eficiência poderia promover um maior conhecimento de utilidade científica e prática para a definição de quais atividades e condições propiciam o alcance do máximo de resultados do treinamento com a menor utilização possível de recursos.

A limitação acerca de como a avaliação de eficiência é ainda pouco explorada no campo de TD&E pode ser suprida com a incorporação de modelos de avaliação oriundos de outros campos, que a consideram um critério de valor na avaliação (Lagbein & Felbinger, 2006). A perspectiva da eficiência constitui um critério de avaliação tradicionalmente investigado pelo campo da Economia (Lagbein & Felbinger, 2006), cujos modelos de análise subsidiam a avaliação de ações, programas e atividades em variados campos de conhecimento (Emrouznejad & Yang, 2018). O conceito de eficiência no âmbito da Economia vai além da relação insumo-produto trazida por indicadores operacionais unidimensionais (Rosano-Peña, 2008), mas é tratado como um fenômeno complexo multivariado e multideterminado, que reflete uma propriedade latente dos sistemas produtivos relativa à sua capacidade de transformação de insumos em produtos e resultados (Bogetoft & Otto, 2011).

Dessa forma, uma relação direta desse conceito econômico de eficiência pode ser realizada para o desenvolvimento de uma abordagem para a avaliação dos Sistemas de TD&E, que integre, da mesma forma, múltiplos fatores processuais e contextuais. Logo, o presente estudo se justifica na medida em que se propõe a incorporar as abordagens econômicas de avaliação da eficiência para o aperfeiçoamento do arcabouço teórico-metodológico do campo de TD&E por intermédio da operacionalização de uma avaliação da eficiência do treinamento com um método pouco utilizado para esses propósitos, qual seja a Análise Envoltória de Dados.

Sob o ponto de vista das contribuições práticas, conforme Passmore & Velez (2015), ao se considerar possíveis lacunas dos modelos de avaliação face ao que tem sido buscado por profissionais de TD&E nas organizações, menciona-se a busca por abordagens relacionadas à definição de medidas de *benchmark* e eficiência da atividade de treinamento. A revisão realizada por Markoulli, Lee, Byington & Felps (2017) demonstra que unidades de gestão de pessoas ainda mantém uma discussão relevante sobre os custos de suas atividades. Essas são evidências da preocupação das organizações com a avaliação do aproveitamento dos recursos investidos na força de trabalho para seu melhor desempenho na geração de resultados monetários e não-monetários.

Portanto, a possibilidade de geração de uma medida de eficiência do sistema de treinamento, isto é, que abranja o programa de ações educacionais e não apenas ações educacionais específicas, aliada à determinação de melhores práticas – *benchmarks* – pode proporcionar um avanço relevante no suprimento das lacunas apontadas sobre a eficiência do treinamento (Passmore & Velez, 2015). Em complemento, para o contexto do setor público em específico, a abrangência de múltiplos programas de treinamento promovendo comparações entre os sistemas de treinamento das organizações, permite a determinação de como as políticas de desenvolvimento de pessoas estão sendo implementadas e em que medida os recursos públicos investidos nessas atividades promovem o alcance de objetivos almejados pelo governo. O monitoramento do custo-eficácia das ações educacionais e do alcance do maior número de servidores, promovendo o desenvolvimento equânime da força de trabalho, pode ser realizado para fins de contribuição à PNDP (Brasil, 2019a).

2 Referencial Teórico

O presente capítulo visa apresentar os principais conceitos componentes do arcabouço teórico que fundamenta este trabalho. Serão inicialmente abordados o conceito de aprendizagem no contexto de trabalho e o desenvolvimento do campo de estudo denominado Treinamento, Desenvolvimento e Educação (TD&E). Sendo um dos objetivos desse campo avaliar a eficiência, eficácia e a efetividade de ações educacionais, é apresentada uma análise dos principais modelos de avaliação do campo e como esses abrangem tais perspectivas de avaliação. Posteriormente, discute-se como o aprofundamento teórico-metodológico acerca da avaliação de eficiência pode complementar a literatura, de forma que se propõe, para tanto, um modelo teórico como subsídio para a avaliação da eficiência em TD&E. Para a operacionalização das perspectivas trazidas pelo modelo são apresentados os principais métodos de avaliação da eficiência, com destaque para a metodologia denominada Análise Envoltória de Dados – DEA, a qual atualmente possui ampla utilização no campo da Educação. Finalmente, é realizada uma revisão sistemática de literatura que congrega os temas de treinamento e avaliação de eficiência, a fim de apresentar o estado da arte e embasar a seleção de variáveis que irão compor o modelo final adotado.

2.1 Aprendizagem e Treinamento no Contexto Organizacional

A aprendizagem de um indivíduo se caracteriza como um processo complexo, dinâmico e permanente, decorrente das interações cotidianas com o seu ambiente. Os processos de aprendizagem em diferentes contextos resultam na gradual aquisição e retenção de conhecimentos, habilidades e atitudes (CHA) que levam à transferência do aprendizado para uma atividade, ou seja, a sua aplicação na realização de tarefas. A capacidade de mobilizar de maneira integrada esses CHA para o desempenho de determinada atividade denota o conceito mais difundido de competência. As competências de um indivíduo, portanto, consistem na combinação dos CHA necessários para o desempenho bem-sucedido de uma atividade, seja ela simples, como escrever uma carta, ou complexa, como projetar um edifício (Brandão, 2007; Pires et al., 2005).

O desenvolvimento de competências individuais torna-se fundamental em contextos organizacionais, na medida em que o desempenho dos indivíduos se reflete no desempenho e na capacidade das organizações de alcançar seus objetivos de natureza econômica e social. Assim, quando a aprendizagem ocorre em contextos de trabalho, essa se volta para a

aquisição de competências profissionais para o cumprimento de atribuições específicas e superação de desafios por meio da execução das atividades requeridas para determinado resultado organizacional almejado (Freitas & Brandão, 2006; Sampaio, Borges-Andrade & Bonatti, 2019).

Sendo um processo contínuo de interação, parte da aprendizagem de um indivíduo em contextos de trabalho ocorre de maneira natural ou, como definido pela literatura, de maneira informal. Os indivíduos, por exemplo, podem lançar mão de estratégias de observação, replicação e busca ativa de conhecimentos, ou engajar-se em processos sociais de aprendizagem por meio da interação com clientes, fornecedores, colegas e superiores ampliando continuamente suas competências. As organizações, por sua vez, por meio de suas políticas e práticas de trabalho, podem promover a aprendizagem de maneira natural quando fomentam um ambiente de colaboração e inovação, quando estabelecem desenhos de trabalho em que a aprendizagem ocorra entre pares na realização de tarefas, a chamada aprendizagem *on-the-job*, ou por meio de processos espontâneos de mentoria entre trabalhadores experientes e novatos. Assim, o contexto de trabalho é, por si só, um espaço de contínua produção de conhecimento coletivo e de promoção de aprendizagem.

Para além desses processos de aprendizagem informal, no entanto, a organização também pode atuar diretamente por meio de intervenções planejadas para a promoção da aprendizagem de competências profissionais. Uma dessas intervenções é o Treinamento, o qual, devido ao seu caráter planejado e sistemático, é compreendido como um processo de aprendizagem formal, ou induzida (Loiola, Nérís & Bastos, 2006; Mourão, Porto & Puente-Palácios, 2014; Salas et al., 2012, Sampaio, Borges-Andrade & Bonatti, 2019; Tannenbaum & Yukl, 1992).

As primeiras práticas formais de treinamento se consolidam como uma atividade organizacional a partir de 1903, juntamente com os princípios da Administração Científica. Nesse contexto, os treinamentos se tornaram uma das principais ferramentas para a preparação de novos trabalhadores recrutados em grandes contingentes pela indústria e para a melhoria da eficiência e desempenho em seus processos de trabalho. Até meados da década de 1940, a função de treinamento era restrita nas organizações, não sendo considerada ainda como uma função especializada, mas como uma atividade conduzida quase exclusivamente por supervisores no próprio ambiente de trabalho e voltada para a identificação e resolução de lacunas de desempenho dos trabalhadores para melhoria da produtividade (Freitas, Jabbour, & Santos, 2011; Pilati, 2006).

Com a intensificação de contextos de conflito e crises econômicas em âmbito mundial, o treinamento passou a ser visto como uma atividade de grande escala, fundamental para o aumento da produção e recolocação de trabalhadores em setores econômicos estratégicos. Dessa forma, paulatinamente, a reponsabilidade por essa atividade foi sendo atribuída não a supervisores, mas a departamentos específicos com profissionais especializados e dedicados especificamente ao planejamento, implementação e avaliação de programas de capacitação nas organizações (Bell et al., 2017; Pilati, 2006). Nesse ínterim, associações de profissionais de treinamento foram criadas em variados contextos e ramos de trabalho, tendo como uma de suas pioneiras a Sociedade Americana para o Treinamento e Desenvolvimento (*American Society for Training and Development – ASTD*), a qual surgiu em 1943. Tais associações, apoiadas no forte interesse que o Treinamento havia despertado a partir da década de 1940, passaram a promover a consolidação profissional e científica desse campo e a promover a especialização dessa função nas organizações (Vargas & Abbad, 2006).

Ao longo do restante do século XX, diante de ambientes de atuação cada vez mais dinâmicos, caracterizados pelo aumento da competitividade em escala global, rápido avanço tecnológico e maior demanda e complexidade dos contextos socioeconômicos, as organizações modernas passaram a se apoiar cada vez mais no treinamento como forma de promover a qualificação contínua de sua força de trabalho (Ferreira & Abbad, 2014; Kraiger et al., 2004; Meneses & Abbad, 2009; Salas & Cannon-Bowers, 2001).

Com isso, a função de TD&E se expandiu, passando a ser compreendida não apenas como solução de problemas de desempenho individual, mas também como fonte de geração de bem-estar e desenvolvimento de trabalhadores, de benefícios para a economia e para a sociedade e como meio de obtenção, por parte das organizações, de vantagens competitivas e capacidades de inovação, sobrevivência e adaptação a mudanças ambientais (Aguinis & Kraiger, 2009; Pilati, 2006; Salas et al., 2012; Salas & Cannon-Bowers, 2001). Os departamentos de treinamento em algumas organizações passaram a constituir estruturas com maior abrangência, complexidade e autonomia, tais como as universidades corporativas no setor privado (Pilati, 2006) e as escolas de governo no setor público (OCDE, 2017) voltadas especificamente para a provisão de ações educacionais para o treinamento de indivíduos em contextos de trabalho. O caráter especializado dessas unidades denota a sua responsabilidade por implementar uma gestão profissionalizada e cientificamente

fundamentada do Treinamento para que o grande volume de investimentos realizados nessa atividade alcance os resultados almejados.

Devido a esse crescente caráter estratégico do Treinamento nas organizações, os investimentos nessa atividade aumentaram continuamente de forma relevante. Atualmente, tem sido recorrente na literatura de Treinamento a contextualização relativa ao patamar bilionário dos gastos com treinamento realizados pelas organizações públicas e privadas (e.g. Bell et al., 2017; Garavan et al., 2019; Kennedy et al., 2014; Kraiger et al., 2004; Meneses et al., 2010; Salas & Cannon-Bowers, 2001; Salas et al., 2012; Laval Silva, 2011; Tharenou, Saks & Moore, 2007). Nesse ínterim, conforme explicitado por Meneses et al. (2010), a importância, complexidade e a dimensão que o TD&E adquiriu propiciou a demanda por uma grande quantidade de pesquisas nessa área e a consequente consolidação de um campo científico específico para seu estudo denominado de Treinamento, Desenvolvimento e Educação (TD&E).

Fundado na Psicologia Organizacional e do Trabalho, campo no qual a aprendizagem se constitui objeto de estudo, o TD&E desenvolveu um amplo arcabouço técnico-científico para o planejamento, execução e avaliação dos resultados de ações educacionais e que possibilita a maior efetividade dos investimentos realizados pelas organizações nessa atividade (Abbad et al., 2003; Bell et al., 2017; Borges-Andrade, 2006b; Pilati, 2006; Salas et al., 2012; Salas & Cannon-Bowers, 2001).

A Psicologia Organizacional e do Trabalho tem como objeto de estudo tanto os processos de aprendizagem informal quanto formal. Ambos coexistem na organização, de forma que são compreendidos como processos articulados e complementares, que se integram na contribuição para o desenvolvimento profissional dos trabalhadores (Borges-Andrade & Pagotto, 2010). No entanto, a aprendizagem informal tem uma natureza espontânea, emergente, contínua e descentralizada, enquanto o Treinamento, por sua natureza formal, é estruturado como um conjunto de processos organizacionais, desenvolvidos pelo campo mais específico de TD&E com o objetivo de garantir a eficiência, eficácia e efetividade dos esforços organizacionais de treinamento para o alcance dos resultados em relação aos seus objetivos (Abbad et al., 2020; Borges-Andrade et al., 2006). Nesse sentido, com vistas a delinear os fundamentos paradigmáticos e teóricos que orientaram o desenvolvimento desse referido arcabouço teórico-metodológico, a próxima seção apresenta a evolução do campo de TD&E desde suas origens até a atualidade.

2.1.1 *Evolução do campo de TD&E*

A crescente importância dos treinamentos nas organizações refletiu-se no aumento da produção científica acerca dessa atividade. Conforme revisão de Bell et al. (2017), um dos primeiros artigos científicos sobre esse tema foi publicado ainda em 1918. O estudo trouxe o delineamento de um plano instrucional para a padronização dos treinamentos para profissionais do campo da Psicologia Aplicada, uma profissão que havia crescido em número e importância durante a Primeira Guerra Mundial. A partir de então, um contingente cada vez maior de pesquisas sobre treinamentos passou a ser publicado com foco semelhante na descrição ou avaliação de treinamentos específicos nos setores público e privado.

Nas décadas seguintes, a aprendizagem tornou-se objeto de estudos aprofundados com pesquisas descrevendo os elementos envolvidos na aquisição de diferentes habilidades e investigando níveis de aprendizagem (Bell et al., 2017). Com um aumento dramático no número de artigos publicados, a partir de 1940 a literatura se voltou para a identificação dos fatores preditivos do sucesso dos treinamentos e, em 1950, um dos primeiros estudos sobre ensino à distância é publicado em função de novos adventos tecnológicos, como a televisão. Segundo Bell et al. (2017), também surgiram nesses períodos os primeiros artigos com foco na definição de métodos e instrumentos para avaliação de treinamentos acerca dos seus resultados de aprendizagem e transferência de conhecimentos.

Grande parte das pesquisas de aprendizagem aplicada nesse período foi apoiada por agências de pesquisa militar (Bell et al., 2007), possivelmente em razão do contexto legado pelas Guerras Mundiais. Nesse contexto, a partir da década de 1960 os estudos em organizações militares realizados por Gagné (1970), denominados “princípios da aprendizagem” (Campbell, 1971, p. 566; Meneses et al., 2010, p. 28) estabeleceram um marco fundamental para o campo de TD&E, assentando-o firmemente sobre a emergente perspectiva da Teoria Geral dos Sistemas (TGS) e que vieram a propiciar a organização e evolução do TD&E como um campo científico específico nas décadas seguintes (Bell et al., 2017; Pilati, 2006;).

A TGS havia sido desenvolvida na década de 1950 no campo da matemática e postulava, de forma geral, a descrição de fenômenos por meio de uma perspectiva cibernética, nas quais seus elementos poderiam ser considerados fundamentalmente como sistemas constituídos por um fluxo de entradas (*input*), processos, saídas (*output*) e retroalimentação (*feedback*), em constante interação com seus ambientes. Tal perspectiva terminou por se constituir como uma base paradigmática comum para variados campos

científicos ao possibilitar a concepção de seus fenômenos como sistemas abertos, complexos e integrados, em oposição à predominante perspectiva mecanicista da ciência (Rousseau, Wilby, Billingham, & Blachfellner, 2018; Scott, 1998). A partir da década de 1960, os pressupostos da TGS foram gradualmente incorporados pelas Ciências Organizacionais como ferramentas de diagnóstico de sistemas organizacionais enquanto estruturas de processamento de informação e tomada de decisão (Kast & Rosenzweig, 1972; Katz & Kahn, 1978; Scott, 1998).

Por meio da perspectiva trazida por Gagné (1970), o campo de TD&E também se tornou pioneiro nessa transição paradigmática. Partindo da concepção do Treinamento como um sistema aberto, Gagné incorpora uma característica fundamental dos sistemas abertos, qual seja a sua orientação para uma finalidade, propósito ou objetivo (Katz & Kahn, 1978). No contexto organizacional, tais objetivos constituem os resultados almejados que possuem valor para o ambiente e para o próprio sistema, alcançados por meio de suas atividades (Scott, 1998). Nesse ínterim, a perspectiva sistêmica do TD&E parte da compreensão prévia dos objetivos e necessidades dos indivíduos e da organização, para que, com o conhecimento aprofundado acerca do trabalho a ser executado, realize um planejamento efetivo para suprimento das necessidades de aprendizagem detectadas com a finalidade última de promover o desempenho e resultados previstos para as atividades organizacionais (Campbell, 1971; Pilati, 2006).

Segundo Bell et al. (2017), tais avanços nas décadas de 1960 e 1970 estabeleceram a concepção do TD&E como um campo científico, evidenciada, por exemplo, pela publicação da primeira revisão integrativa de literatura específica deste campo por Campbell (1971). Nesse artigo, Campbell (1971) revisou os principais avanços do campo de TD&E, delineando os primeiros estudos acerca de seus componentes, assim como as principais teorias, perspectivas e objetos de estudo relacionados à aprendizagem, ao desenho instrucional e à avaliação do treinamento, os quais viriam a se tornar grandes linhas de desenvolvimento científico. A revisão de Campbell (1971) demonstrou os numerosos estudos no campo e, em tom crítico, exortou os pesquisadores a irem além, na concretização de uma base empírica e metodológica robusta para o campo.

Talvez em razão desse chamado de Campbell (1971), o campo continuou a crescer rapidamente em volume e escopo, destacando-se a inclusão de metodologias científicas de desenho experimental, avaliações comparativas de treinamentos, análise de efeitos do treinamento no ambiente organizacional e, a partir da década de 1980, a análise das

influências de características e percepções de indivíduos na dinâmica e nos resultados da aprendizagem, em perspectivas cognitiva e comportamental. Além dos fatores individuais, a década de 1990 é marcada pela compreensão da influência dos fatores contextuais na aprendizagem antes, durante e após o treinamento, tais como o suporte organizacional e os desenhos de trabalho. Tais perspectivas, que passam a considerar características individuais e fatores contextuais, dão origem a uma renovação dos paradigmas e critérios para a avaliação dos treinamentos com a consolidação de modelos sistêmicos que incorporam essas variáveis e levam à maior compreensão do papel da organização e dos gestores na geração das condições necessárias para o sucesso do processo de aprendizagem (Bell et al., 2017).

Por fim, Bell et al. (2017) apresentam como os avanços a partir dos anos 2000 fornecem uma compreensão ampliada do Treinamento como um processo que tem o condão de aperfeiçoar não apenas indivíduos e organizações, mas o contexto social como um todo (Aguinis & Kraiger, 2009). Nesse período, além da expansão dos estudos acerca dos fatores individuais e contextuais que afetam o treinamento, a perspectiva instrucional foi revolucionada com estudos sobre as emergentes tecnologias de informação e comunicação e suas implicações para os treinamentos, de forma que múltiplas técnicas instrucionais inovadoras são estudadas.

As revisões e a definição de diretrizes para o campo de TD&E, iniciadas por Campbell (1971), foram continuadas, por exemplo, por Goldstein (1980), Wexley (1984), Latham (1988), Tannenbaum & Yukl (1992), Salas & Cannon Bowers (2001), Aguinis & Kraiger (2009), Salas et al. (2012), pelo próprio trabalho de Bell et al. (2017) e mais recentemente pela revisão de Kraiger & Ford (2021). Tais revisões evidenciam a consistência, consolidação e a unicidade teórica do campo de TD&E em torno de seus componentes fundamentais e documentam os avanços empíricos acerca de seus principais fatores interferentes, assim como da eficácia e efetividade de variados tipos de treinamento em diferentes contextos para a aprendizagem de indivíduos nas organizações.

Assim, ao longo de mais de um século desde seu surgimento, o campo de TD&E tem se desenvolvido de maneira consistente tornando-se, atualmente, um dos temas mais abrangidos no âmbito das literaturas de Psicologia e Administração, tanto no contexto internacional (Bell et al., 2017; Garavan et al., 2019; Markoulli et al., 2017; Salas & Cannon-Bowers, 2001;) quanto no contexto nacional de pesquisa (Borges-Andrade & Pagotto, 2010; Gondim et al., 2019). Considerado por muitos já como uma ciência em si, o campo de TD&E estabeleceu extensiva literatura científica e prescritiva para estudo e aplicação no contexto

das organizações (Abbad et al., 2003; Bell et al., 2017; Borges-Andrade, 2006a; Pilati, 2006; Salas et al., 2012; Salas & Cannon-Bowers, 2001).

Conclusivamente, como pôde-se observar, o presente capítulo partiu da descrição do Treinamento enquanto atividade organizacional, destinando-se à descrição do campo científico de TD&E. Os três conceitos implicados na denominação do campo refletem uma evolução conceitual e uma importante distinção entre Treinamento, Desenvolvimento e Educação, fundamental para a compreensão do objeto de estudo do campo. Portanto, de forma complementar à presente revisão histórica do TD&E, a próxima seção visa delinear um recorte acerca de como a evolução do campo refletiu-se em uma gradual modificação de seu conceito e distinções entre as ações educacionais realizadas nas organizações.

2.1.2 O conceito de Treinamento em TD&E

O campo de estudo voltado para a aprendizagem nas organizações denominou-se inicialmente “Treinamento”, sendo conceituado em termos de seus aspectos fundamentais como um esforço formal de capacitação para os contextos de trabalho. Dessa forma, fundamentalmente, Treinamento é definido como a aquisição e desenvolvimento sistemáticos de conhecimentos, habilidades e atitudes requeridos por empregados para execução adequada e melhoria de desempenho de tarefas no trabalho (Goldstein, 1980; Latham, 1988). O termo sistemático remete ao aspecto formal da aprendizagem, caracterizando-se como um processo técnico e gerencial que executa uma série de atividades planejadas pela organização para promover a aprendizagem pelos empregados de comportamentos relacionados ao trabalho (Wexley, 1984).

No entanto, ao longo do desenvolvimento do campo, sobretudo com o surgimento da *American Society for Training and Development* – ASTD, o termo “Desenvolvimento” passou a integrar o campo que foi então denominado “Treinamento e Desenvolvimento” (T&D). O novo acrônimo, que vigora até a atualidade no contexto internacional, visava integrar ao campo conceitos mais amplos de desenvolvimento de recursos humanos em perspectivas de aprendizagem não necessariamente vinculadas ao trabalho (Vargas & Abbad, 2006).

Para Garavan (1997), tal mudança não ocorreu sem que houvesse uma dificuldade em se realizar uma distinção teórica precisa entre os dois termos. Em razão dessa dissonância, parte dos teóricos do campo concebeu o “Desenvolvimento” como um termo que já implicava o “Treinamento”, de tal forma que um novo campo científico fora proposto,

especificamente por Nadler (1984), o qual denominou-se “Desenvolvimento de Recursos Humanos” (DRH) (Stewart et al, 2010). O DRH surge com a proposta de integrar teoricamente a perspectiva de estudo micro organizacional centrada no indivíduo, com a perspectiva macro organizacional, centrada na organização, e de expandir seu escopo para além das teorias de aprendizagem (Stewart et al, 2010). Segundo Garavan, Heraty & Barnacle (1999), no entanto, a sobreposição do DRH com outros campos, como o de TD&E e o de Gestão de Pessoas, levou a uma fragmentação problemática, constituindo uma literatura que não é homogênea nem claramente definida. Não obstante, o campo de DRH tem se desenvolvido, mas em recente revisão, Garavan et al. (2019) exploram com mais clareza e expõem de maneira crítica as lacunas teórico-metodológicas enfrentadas pelo campo em sua busca por um elo entre aprendizagem e desempenho no nível individual e os resultados no nível organizacional.

Por outro lado, a corrente que se manteve atinente ao conceito de T&D, tendo em vista os fundamentos do campo na Psicologia Organizacional, manteve como escopo fundamental a aprendizagem nas organizações. Para o campo de T&D, a aprendizagem é um fenômeno transversal tanto ao “Treinamento” quanto ao “Desenvolvimento” (Garavan, 1997), de forma que se manteve a unicidade epistemológica do campo, solucionada por meio de uma distinção conceitual entre esses termos que seriam, de toda maneira, englobados pelas teorias do campo como um todo. No Brasil, especificamente, o termo “Educação” foi incorporado, constituindo-se, assim, a denominação do TD&E. Segundo Vargas & Abbad (2006), essa expansão traduz a compreensão da aprendizagem como um fenômeno de longo prazo nas organizações, de forma que deve se integrar com o desenvolvimento e a ascensão profissional dos indivíduos na organização. De forma geral, a denominação “TD&E” implica uma transição da simples preocupação com o desempenho eficiente e eficaz no posto de trabalho para o envolvimento organizacional com o desenvolvimento humano de seus empregados e o reconhecimento de seus anseios e aspirações (Vargas & Abbad, 2006).

Consoante ao exposto por Garavan (1997), Abbad e Borges-Andrade (2014) explicam que os termos Treinamento, Desenvolvimento e Educação muitas vezes são utilizados equivocadamente como sinônimos, embora sejam referentes a distintas e igualmente importantes atividades de apoio à aprendizagem. A literatura nacional, ao longo de amplo debate, buscou estabelecer uma distinção entre os termos que, em suas proposições iniciais, baseava-se na finalidade e propósito da ação educacional empreendida (Meneses et al., 2010).

Sob essa perspectiva, o Treinamento é compreendido como uma ação educacional destinada ao aperfeiçoamento de um indivíduo para um desempenho específico em suas atividades atuais de trabalho visando objetivos de desempenho específico. O Desenvolvimento, por sua vez, consiste em ações planejadas e direcionadas para o crescimento pessoal do empregado, sem manter relações estritas com o trabalho em sua organização. Por fim, a Educação refere-se às oportunidades dadas pela organização ao indivíduo visando prepará-lo para ocupar novos cargos na organização, em uma perspectiva orientada para o longo prazo. (Meneses et al., 2010; Vargas & Abbad, 2006).

Observou-se, no entanto, que tal distinção poderia ser relativizada, tendo em vista que uma ação educacional poderia compreender múltiplas finalidades. Em suma, não se poderia definir categoricamente que ações de Educação e Desenvolvimento não gerassem CHA úteis para o desempenho no cargo atual, tampouco que o Treinamento não proporcionasse o crescimento pessoal ou oportunidades de ascensão profissional. Por esse motivo, Vargas & Abbad (2006) propõem não mais a distinção desses termos em razão apenas da sua finalidade, mas em razão principalmente “da complexidade da estrutura do conhecimento das ações educacionais” (Meneses et al., 2010, p. 19). As proposições de Vargas & Abbad (2006) são sumarizadas, portanto na Figura 1, na qual é possível compreender como se organizam esses conceitos e as respectivas ações educacionais, distintas em complexidade, que lhes são associadas de forma exemplificativa.

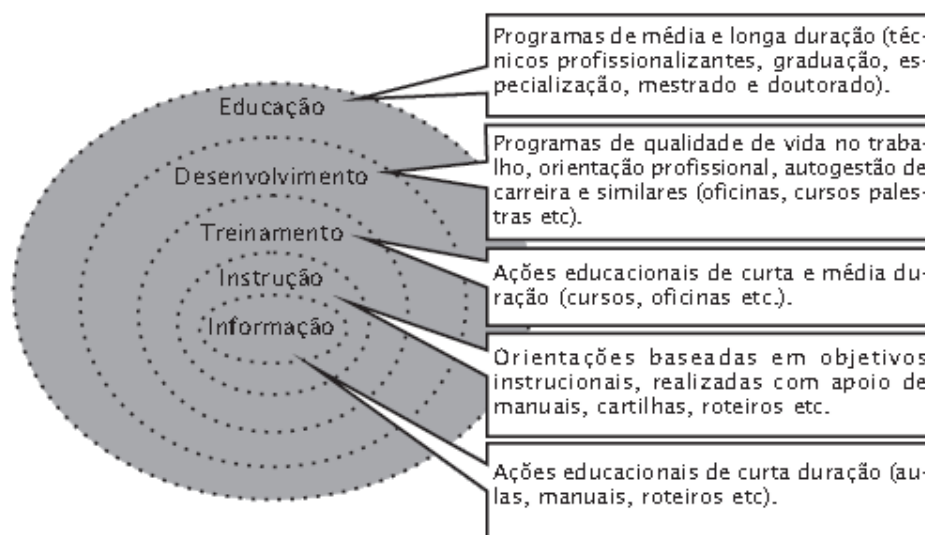


Figura 1. Relação entre conceitos e respectivas soluções educacionais

Fonte: Vargas e Abbad (2006)

Como visto na Figura 1, o mais amplo dos termos concerne à “Educação”, que trata de uma forma ampla de aprendizagem de longo prazo que vai além de um contexto específico e imediato de trabalho, visando preparar o indivíduo para o futuro. Baseando-se em Nadler (1984), Vargas & Abbad (2006) estabelecem a educação como mais ampla e complexa na medida em que esta tem assumido novos significados, incorporando tanto o crescimento profissional quanto pessoal (Meneses et al., 2010). As ações mais adequadas para representar o conceito de Educação se referem a cursos de média e longa duração, tais como, cursos técnicos profissionalizantes, de graduação e pós-graduação.

As ações de “Desenvolvimento”, na medida em que não se relacionam necessariamente ao desempenho de tarefas de trabalho (Vargas & Abbad, 2006), estão relacionadas a cursos e palestras com o objetivo de promover o crescimento pessoal do indivíduo em níveis de complexidade mais simples do que as ações associadas à Educação (Meneses et al., 2010). O Treinamento tem sua menor complexidade definida pela especificidade dos conhecimentos, mais restritos ao desempenho em uma atividade específica e alinhados aos desempenhos requeridos pela organização (Vargas & Abbad, 2006). Por fim, em linha com o critério de complexidade adotado pelas autoras, os formatos representados de Informação e Instrução podem ser considerados formas mais simples e padronizadas de alcance desses objetivos, com a disponibilização de informações relativas ao trabalho para induzir processos de aprendizagem, em geral, de maneira individual (Meneses et al., 2010).

Dessa forma, Treinamento, Desenvolvimento e Educação são compreendidos como tipos de ações de aprendizagem realizadas por uma organização, porém com distintos níveis de complexidade. Considerando a perspectiva sistêmica do TD&E, a diferença entre essas ações com relação à complexidade para sua estruturação e execução, assim como as diferentes finalidades que podem atender, podem implicar em diferentes fluxos ou mesmo diferentes sistemas de aprendizagem em uma organização.

Consoante ao proposto por Becker & Huselid (2006), as organizações podem desenvolver políticas e práticas de gestão de pessoas segmentadas – ou seja, múltiplos sistemas de Gestão de Pessoas com diferentes filosofias, processos e regras – para, por exemplo, diferentes grupos de profissionais (Becker & Huselid, 2006). Tal asserção também é proposta por Gratton (1999) quando distingue um sistema de desenvolvimento de pessoal orientado para o curto prazo e outro orientado para o longo prazo.

Consequentemente, é possível que o sistema de TD&E deva empregar abordagens distintas para avaliação de necessidades, planejamento, execução e sobretudo avaliação de ações de aprendizagem, conforme o tipo de ação educacional. Como exposto por Abbad et al. (2020) nem todas as ações educacionais se destinam a gerar resultados organizacionais. Tomando por exemplo as ações de “Desenvolvimento”, dada sua finalidade mais ampla, pode não ser adequada a busca pela mensuração de seus resultados. De forma similar, os resultados de ações Educacionais, projetados em longo prazo, implicam em formas de avaliação que podem ser distintas daquelas adotadas para o Treinamento, orientado para resultados específicos e de curto prazo. Tendo em vista as diferenças entre Treinamento, Desenvolvimento e Educação, a consideração dessas ações como similares, como alertado por Garavan (1997) e Abbad e Borges-Andrade (2014), pode levar a equívocos tanto em contextos de pesquisa quanto aplicados.

Portanto, para os fins de análise das sistemáticas de avaliação em TD&E, tal como a intentada na presente pesquisa, a complexidade estrutural de ações compreendidas como esforços de Educação podem ensejar o cuidado de trata-las de maneira distinta de ações menos complexas como as de Treinamento e Desenvolvimento. O estudo da Educação tem ampla literatura específica (e.g. De Witte & López-Torres, 2017), de forma que seu escopo, de certa maneira, se distingue daquele abrangido por práticas de Treinamento e Desenvolvimento, ainda que haja sobreposição com relação às suas finalidades e variáveis de estudo. Nesse ínterim, para os propósitos dessa pesquisa o escopo específico sobre as quais se debruçarão as principais análises corresponderá às ações compatíveis com os níveis de complexidade do Treinamento e Desenvolvimento, abrangendo de forma mais enfática o conceito de Treinamento, haja vista que se relaciona mais diretamente à finalidade e à premissa predominante de contribuição aos objetivos organizacionais.

A despeito dessas distinções, a maior parte das ações educacionais desenvolvidas em âmbito organizacional é passível de um processo sistemático de planejamento, execução e avaliação (Salas et al., 2012). As atividades desse processo podem ser representadas por um modelo para o Sistema de Treinamento, definido pelo campo de TD&E como um conjunto cíclico de práticas interligadas que visa promover o desenvolvimento de pessoas de forma alinhada aos objetivos organizacionais (Bell et al, 2017; Salas et al, 2012). Nesse sentido, a próxima seção descreve detalhadamente os componentes desse sistema para que se defina com precisão como o TD&E busca promover a eficiência, eficácia e a efetividade dos investimentos em ações educacionais.

2.1.3 O Sistema de TD&E

O TD&E é concebido a partir de uma premissa sistêmica, sendo compreendido dessa forma como um conjunto de processos integrados, realizados no contexto organizacional, com o qual mantém uma relação de interdependência (Borges-Andrade et al., 2012; 2013; Goldstein, 1990; Meneses et al., 2010; Pilati, 2006). Convencionalmente, três processos constituem o sistema de TD&E, conforme explicitado pela Figura 2: a avaliação de necessidades de treinamento (ANT); planejamento e execução; e a avaliação de treinamento (Borges-Andrade & Abbad, 1996; Tharenou, 2010).

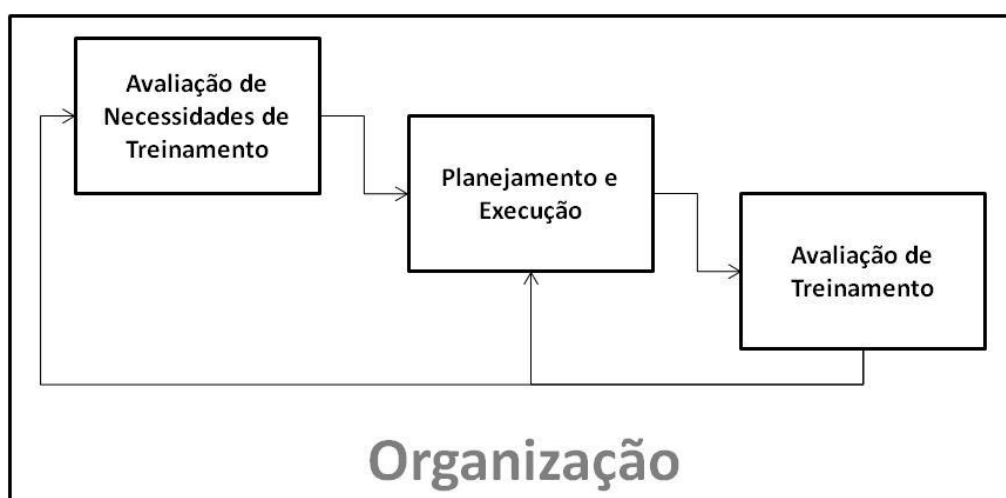


Figura 2. Sistema de Treinamento

Fonte: Borges-Andrade & Abbad (1996)

O primeiro componente do sistema de treinamento, a ANT, consiste no levantamento sistemático de informações para determinação de onde os treinamentos devem ser conduzidos na organização (análise organizacional), o que deve ser treinado em termos de conhecimentos, habilidades e atitudes requeridos pelo trabalho (análise de tarefas) e quais os indivíduos que necessitam de determinados treinamentos (análise individual) (Goldstein, 1980; Latham, 1988; Meneses et al., 2010; Pilati, 2006; Salas et al., 2012; Tharenou, 2010; Wexley, 1984).

A Avaliação de Necessidades possui um caráter descritivo e diagnóstico (Lagbein & Felbinger, 2006) devendo orientar-se para a detecção de lacunas de desempenho individual e organizacional discernindo aquelas que podem ser solucionadas por intermédio de intervenções de aprendizagem (Chyung, 2019). Com isso, é fundamental que a ANT resulte em um plano a ser implementado pela organização orientado para objetivos de aprendizagem

a serem supridos por um programa de ações educacionais que sejam coerentes e priorizadas conforme as lacunas de desempenho levantadas para os indivíduos e para a organização. Segundo Chyung (2019), tais planos também devem envolver aspectos de implementação, avaliando os meios necessários e estratégias mais adequadas para a solução dos problemas organizacionais. Para Ferreira & Abbad (2004), também devem resultar da ANT os parâmetros para a avaliação dos treinamentos realizados frente aos objetivos definidos.

O segundo componente do sistema de treinamento refere-se ao planejamento e execução dos cursos pela organização. Por meio da aplicação de princípios de psicologia instrucional, são planejadas as ações educacionais e a melhor forma de articulação de objetivos de aprendizagem, métodos, estratégias, ferramentas e recursos instrucionais visando facilitar o processo de ensino-aprendizagem. Nessa etapa é que se define qual o melhor tipo de ação educacional, consideram-se alternativas de implementação, com relação, por exemplo, à sua modalidade de ensino, estrutura, conteúdos e recursos humanos, financeiros e materiais requeridos e o perfil dos participantes. Tal processo, haja vista a variedade de ações educacionais, é realizado continuamente e deve ter sempre por referência a melhor forma de promover o suprimento das necessidades de desenvolvimento definidas na etapa de ANT de maneira eficiente e eficaz (Borges-Andrade, 2006a; Meneses et al., 2010; Pilati, 2006; Salas et al., 2012).

Por fim, o terceiro componente do sistema de treinamento consiste na avaliação das ações educacionais, a qual mantém uma relação recursiva com os demais subprocessos de treinamento, haja vista que considera como critério de avaliação os objetivos pré-estabelecidos nessas etapas anteriores. Seu principal objetivo é avaliar o sucesso dos treinamentos na modificação de comportamentos e aperfeiçoamento de desempenho dos indivíduos, buscando estabelecer a efetividade do Treinamento (Abbad et al., 2020; Passmore & Velez, 2015; Tharenou, 2010).

Cabe destacar que conforme Borges-Andrade (2006) e Salas et al. (2012), além dos processos fundamentais do Sistema de Treinamento, há uma série de atividades que devem ser empreendidas pelos gestores de TD&E. Para esses autores, dentre as medidas na etapa pré-treinamento, além de uma rigorosa ANT, deve se promover um clima de aprendizagem no trabalho por meio uma estratégia efetiva de comunicação das políticas de TD&E, divulgação adequada das ações ofertadas e a sensibilização das chefias para que apoiem a aprendizagem e a transferência do treinamento. Durante o treinamento, as estratégias instrucionais devem promover a motivação, auto eficácia e a associação dos conhecimentos

aprendidos com a realidade de trabalho dos participantes de treinamento. Por fim, o papel do TD&E não se encerra com as ações educacionais, devendo atuar em uma etapa pós-treinamento de acompanhamento (*follow-up*) para remover obstáculos, prover orientações e reforços junto aos indivíduos e suas chefias para que se assegure a transferência do conhecimento para o trabalho. Por fim, conforme mencionado, após o treinamento procede-se a uma definição clara de quais avaliações devem ser realizadas abrangendo, se possível, diferentes tipos de resultados a fim de atestar o alcance dos objetivos para os indivíduos e para a organização (Abbad et al., 2020).

Assim, na medida em que desenvolveram uma perspectiva sistêmica, a partir de 1970, variados modelos de avaliação foram desenvolvidos para que se ampliasse, gradualmente, os tipos de resultados analisados e os fatores condicionantes do sucesso dos treinamentos. Parte desses modelos incorpora pressupostos de teorias de base do campo da Avaliação e da Educação, tais como o modelo Contexto-Insumo-Processo-Produto (CIPP) de Stufflebeam (2003). Os fundamentos teóricos do campo de Avaliação em específico, permitem uma compreensão aprofundada das características dos tipos e modelos de avaliação em TD&E, os quais têm crescido em quantidade e escopo nos últimos anos (Passmore & Velez, 2015). Dessa forma, o próximo capítulo é voltado para a discussão dos variados objetivos e perspectivas de análise dos processos avaliativos de forma geral e como o campo de TD&E incorporou e operacionalizou tais pressupostos nas atividades específicas de Treinamento organizacional.

2.2 Fundamentos e Perspectivas da Avaliação de Programas

Para Shaw et al. (2006), a avaliação é, em princípio, um processo natural presente em toda atividade humana em que as pessoas utilizam constantemente as informações de que dispõem para realizar julgamentos subjetivos acerca do valor de ações, pessoas, objetos, organizações e situações cotidianamente. Na medida em que julgamentos individuais não explicitam quais os critérios utilizados e constituem um modo informal e subjetivo de avaliação, a realização de avaliações como meio de aperfeiçoamento de sistemas sociais passou a requerer a construção de modelos formais e sistemáticos de avaliação que explicitam os objetivos e os critérios utilizados para atribuição de valor às atividades humanas (Mark et al., 2000; McLaughlin & Jordan, 2015).

As avaliações sistemáticas são utilizadas com variados propósitos, dentre eles a formulação de políticas públicas, melhoria de programas organizacionais, desenvolvimento

de conhecimentos, avanço da justiça social e julgamento e aprimoramento de boas práticas (Mark et al., 2000). Segundo os autores, a natureza difusa e natural do processo avaliativo não permite atribuir-lhe uma única origem enquanto um campo de estudo das ciências sociais. No entanto, parte da literatura de Avaliação tende a apontar o pioneirismo dos trabalhos de Ralph W. Tyler, que estabeleceu, em 1930, uma distinção entre a simples mensuração de dados e a realização de avaliações, em que a avaliação se orienta para a consecução de objetivos. Seu trabalho pioneiro realizado no campo da Educação, visava determinar os efeitos de inovações implementadas por educadores em instituições de ensino (Kellaghan, Stufflebeam & Wingate, 2003; Mark et al., 2000; Stufflebeam & Coryn, 2014).

Como um processo formal e sistemático, Scriven (1991) define que a avaliação pode ser compreendida como uma ferramenta que se caracteriza por um processo sistemático de coleta de dados para determinação do mérito ou do valor de algo. Com esse propósito, ao longo do século XX, as teorias de avaliação se consolidam e se difundem, de tal forma que a Avaliação se torna um campo teórico-metodológico transdisciplinar que fundamenta processos avaliativos em múltiplos campos de conhecimento, sobretudo das ciências sociais aplicadas, onde constituiu-se o subcampo de avaliação de programas (Scriven, 1991; Stufflebeam, 2003).

Segundo McLaughlin & Jordan (2015, p.7) “um programa é um conjunto de recursos e atividades direcionados a um ou mais objetivos comuns, normalmente sob a direção de um único gestor ou equipe de gestão”. Sendo uma definição ampla, um programa pode abranger atividades em qualquer contexto organizacional e com distintos níveis de complexidade. Porém, um programa se caracteriza por almejar objetivos claros, normalmente voltado para a resolução de um problema relevante ou para a promoção de mudanças nos comportamentos dos indivíduos ou do ambiente (Suchman, 1967; McLaughlin & Jordan, 2015). A avaliação de um programa, portanto, consiste na aplicação de métodos sistemáticos para definição do mérito e valor sobre suas operações e resultados (McLaughlin & Jordan, 2015).

Na medida em que parte de uma definição apriorística do que seria o mérito e o valor do programa, e que podem haver distintas perspectivas sobre o que será considerado como um resultado bem-sucedido, as avaliações terminam por incorporar percepções sociais e culturais compartilhadas, de forma que não possuem um caráter inequívoco e uniforme para a representação da realidade (Mark et al., 2000). Segundo Stufflebeam & Coryn (2014) as avaliações podem, portanto, ter uma parcela de subjetividade. Além disso, uma avaliação busca considerar todos os critérios relacionados à determinação de valor, o que a torna

relativamente distinta da pesquisa científica, que se voltará para problemas de pesquisa e hipóteses específicas, ensejando a restrição e o controle de variáveis. Não obstante, a partir do momento em que um consenso acerca dos critérios de valor é estabelecido, o papel do processo avaliativo é torna-lo o mais objetivo, livre de vieses e neutro quanto possível, por meio, portanto, da definição de um método sistemático e objetivo de análise (Suchman, 1967), podendo envolver desenhos metodológicos de base científica para o teste de teorias e relações subjacentes aos critérios de avaliação.

Segundo Stufflebeam e Coryn (2014), Suchman (1967) foi um dos pioneiros na proposição da chamada “pesquisa avaliativa”. Tendo em vista que diferentes formas de avaliação com diferentes graus de subjetividade podem ser realizadas (Stufflebeam & Coryn, 2014) – por exemplo, um júri – Suchman (1967) estabelece que uma das formas de se garantir um nível adequado de neutralidade, rigor e objetividade é buscar uma abordagem científica para a avaliação. Tal abordagem adotada, portanto, no meio científico, implica, segundo Lagbein & Felbinger (2006), que os critérios de valor definidos para as avaliações, em geral, refletem a natureza e as questões de pesquisa do respectivo campo de conhecimento. Avaliações no campo de públicas, por exemplo, tenderão a analisar a efetividade de programas de governo em termos de resultados para a sociedade, enquanto que estudos em Economia e Finanças priorizarão aspectos como custo-benefício, produtividade, lucratividade ou desempenho de mercado.

Com isso, como observa Suchman (1967), certo debate emergiu entre as ciências sociais acerca de quais critérios de valor seriam mais apropriados. Visando integrar as distintas perspectivas de análise, o autor postulou que não há uma melhor e única forma de avaliação, mas distintas categorias igualmente importantes e que denotam dimensões para os objetivos e critérios sobre a qualidade e sucesso de um mesmo programa ou atividade organizacional. Para tanto, propôs cinco categorias que representam diferentes critérios de sucesso (ou fracasso) pelos quais um programa pode ser julgado: (1) Esforço, (2) Desempenho, (3) Adequação do Desempenho, (4) Eficiência e (5) Processo. Tais categorias são descritas a seguir:

1) **Esforço:** As avaliações nesta categoria têm como critérios de sucesso medidas da quantidade e qualidade das atividades realizada no programa. Nesse sentido, relaciona-se aos meios, e aos insumos apenas visando indicar o que está sendo feito e se tais esforços e recursos são suficientes para o alcance dos objetivos de um programa. Tal

avaliação não se relaciona a produtos e resultados, sendo um critério necessário, mas não suficiente para atestar o sucesso do programa.

2) **Desempenho:** As avaliações de desempenho, por outro lado, adotam como critérios os resultados do esforço. Para tanto, é necessária a definição clara de quais eram os objetivos almejados em diferentes níveis, desde os mais imediatos até os efeitos finais, sendo que a sua mensuração deve atender a critérios de validade, ou seja, corresponder a uma medida livre de vieses que reflita concretamente o desempenho realizado.

3) **Adequação de desempenho:** Esta categoria se refere ao grau em que o desempenho realizado, ou seja, as atividades empreendidas, ainda que tenham sido bem-sucedidas, são efetivas com relação à solução do problema ou suprimento das necessidades inicialmente identificadas. Suchman (1967) aponta que medidas de adequação do desempenho consistem, por exemplo, na mensuração do impacto do programa no contexto como um todo e se este atende às expectativas sobre o efeito desejado.

4) **Eficiência:** A avaliação nesta categoria se volta para a consideração de meios ou métodos alternativos, em termos de custos – recursos financeiros, tempo, pessoal e conveniência – para programas e atividades consideradas eficazes e efetivas. A questão que se busca avaliar é se há uma melhor forma de se alcançar os mesmos resultados. A eficiência visa representar “qual a capacidade de um indivíduo, organização, unidade ou atividade de produzir resultados em proporção ao esforço que despende” (Stufflebeam & Coryn, 2014, p. 262). Segundo Suchman (1967), na medida em que poucos programas conseguem justificar completamente os seus custos, medidas de eficiência devem ser incorporadas sempre que possível. Visando justificar a continuidade dos investimentos realizados – já na década 1960 – a competição por fundos entre os diversos serviços públicos ensejou uma grande ênfase em critérios de eficiência. Essa ênfase também se relaciona às revisões promovidas pelos Governos na tentativa de avaliar a continuidade, necessidades de aperfeiçoamento ou a simplificação de programas públicos que podem já ter alcançado os resultados desejados.

5) **Processos:** O objetivo desta categoria é investigar as razões básicas subjacentes a porque um programa teve ou não sucesso no alcance de seus objetivos. Para Suchman (1967), de forma estrita, a análise do processo não faz parte da avaliação em si, na medida em que não se volta para a determinação de seu mérito ou valor, mas a uma consideração de questões de natureza gerencial ou científica para identificar falhas e melhorias, particularmente quando um programa não é bem-sucedido, visando ajustes e

correções para sua continuidade (Stufflebeam & Coryn, 2014). Suchman (1967) delineou quatro dimensões de uma análise de processo: os atributos do programa; a população exposta ao programa; o contexto em que o programa ocorreu; e os diferentes tipos de efeitos produzidos pelo programa. A primeira dimensão visa analisar em detalhe os componentes do programa para encontrar gargalos que impactam toda a atividade. A segunda dimensão considera as características do público-alvo e se este tem sido delimitado, abordado e abrangido de maneira adequada para os efeitos almejados. A terceira dimensão considera possíveis condições contextuais que podem estar afetando o programa. Por fim, a última dimensão refere-se à análise que verifica se os resultados e efeitos do programa são delineados e mensurados de maneira válida e coerente e se decorrem, de fato, das atividades empreendidas.

Segundo Stufflebeam e Coryn (2014), a categorização de Suchman (1967) é considerada uma estruturação pioneira para a definição dos processos de avaliação com a integração de distintas perspectivas de análise abrangendo a totalidade de um programa. Segundo Mark et al. (2000), no entanto, há possivelmente aspectos que podem não se enquadrar nessa categorização, devendo a mesma ser considerada um modelo, mas não uma descrição definitiva sobre as dimensões e critérios das avaliações. O objetivo de Suchman (1967), no entanto, era demonstrar que ao invés de um critério de avaliação, programas públicos e atividades organizacionais deveriam ser abrangidos por um conjunto integrado de avaliações sistemáticas com diferentes critérios e análises para a tomada de decisão, desde a adequação dos insumos (esforço), a definição dos desempenhos necessários, a adequação de seus resultados, a qualidade do processo e, por fim, as alternativas mais eficientes de execução.

Um aspecto de grande relevância trazido por Suchman (1967) é a postulação de uma relação entre essas categorias. Segundo o autor, a inter-relação entre as categorias implica que avaliações de esforço e desempenho necessariamente precedem a avaliação de adequação, eficiência ou processo. Tal relação decorre de uma premissa de encadeamento entre as atividades de um programa em que o esforço bem-sucedido promove o desempenho bem-sucedido em níveis mais amplos de resultado até o alcance do objetivo final (*ultimate objective*). Os programas de avaliação partem, portanto, do conceito de uma cadeia cumulativa de objetivos progredindo do objetivo prático mais imediato em direção ao objetivo final (Suchman, 1967). Uma forma de ilustrar essa lógica de encadeamento é o exemplo genérico trazido pelo autor sobre como um programa de educação em saúde

baseado na distribuição de conteúdo informativo visa alterar costumes de um público-alvo e em última análise, promover a redução de mortalidade (Figura 3).

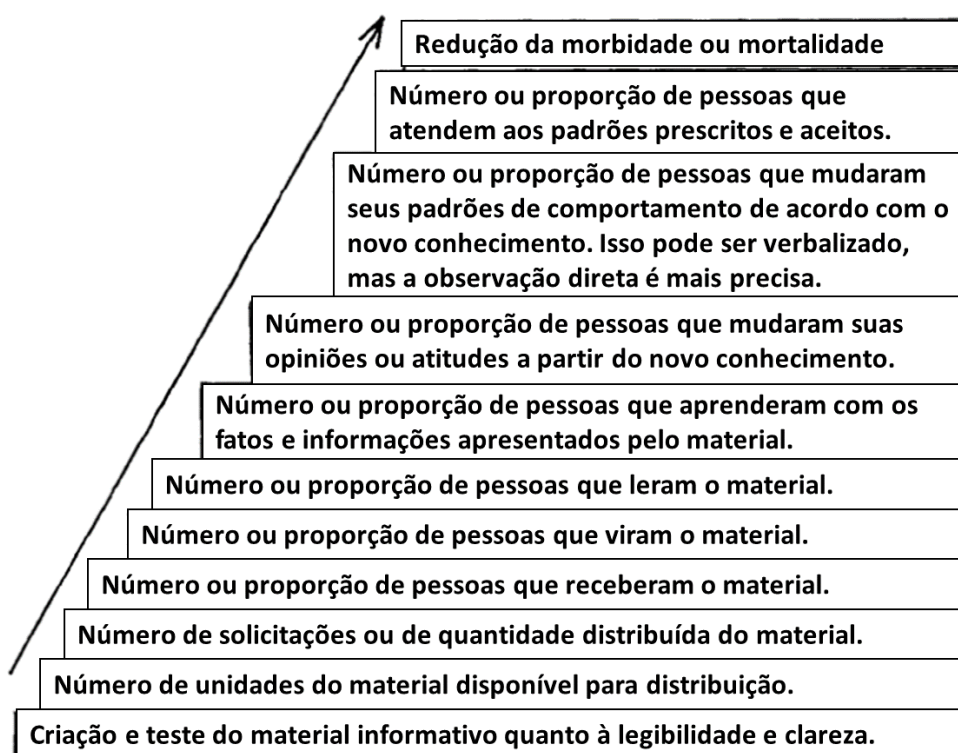


Figura 3. Ilustração da cadeia de objetivos em um programa de educação em saúde
Fonte: Suchman (1967, tradução livre).

A avaliação integrada do processo ilustrado na Figura 3, portanto, demonstra as variáveis que devem ser consideradas pelo avaliador na determinação do sucesso do programa e como elas constituem, em níveis de objetivos, as diferentes perspectivas, que vão desde os esforços práticos, passando por resultados intermediários até os efeitos finais no contexto social mais amplo. Remetendo à necessária abordagem científica subjacente ao processo avaliativo, as relações propostas entre essas variáveis devem ser válidas, fundamentadas na literatura científica ou testadas empiricamente, sendo esse também um dos objetivos da avaliação em perspectiva científica (Suchman, 1967).

Segundo Suchman (1967), a dificuldade dos avaliadores de estabelecer critérios de sucesso advém da confusão entre os níveis, em que determinados avaliadores consideram suficiente o alcance de objetivos intermediários – e.g. aprendizado – enquanto outros consideram como sucesso os efeitos ulteriores de mudança nos comportamentos ou em indicadores sociais. Assim, em realidade, um programa de avaliação idealmente deveria ser capaz de abranger e integrar esses diferentes níveis. No entanto, quanto mais amplo o objetivo avaliado, maior o número de atividades e variáveis que devem ser analisadas e,

portanto, maior a complexidade do processo avaliativo. Dessa forma, deve haver uma consideração do quão factível é a avaliação, assim como das limitações de suas conclusões e capacidade de generalização, na medida em que abranja parcialmente o fenômeno (Suchman, 1967).

Para sistematizar esses diferentes níveis e escopos de avaliação, a literatura tem buscado delinear os tipos e categorias de resultados, os quais convencionalmente, vem sendo compreendidos como de natureza imediata, intermediária e final. Uma “avaliação final” visa determinar o sucesso de um programa em eliminar ou reduzir os problemas identificados. Uma pesquisa “pré-avaliativa” lida com os problemas intermediários que devem ser solucionados anteriormente ao problema final, concentrando-se no diagnóstico das atividades realizadas para aquele fim. Por fim, “avaliações de curto-prazo” limitam-se a responder questões específicas sobre a utilidade e qualidade dos procedimentos realizados (Stufflebeam & Coryn, 2014; Suchman, 1967; McLaughlin & Jordan, 2015).

Os estudos e as proposições de Suchman (1967) estabeleceram bases importantes de valor heurístico, teórico e metodológico para o campo de avaliação e que apoiaram outros importantes teóricos no desenvolvimento das concepções modernas para a compreensão e desenvolvimento de modelos de avaliação. Dentre eles, Scriven (1967 apud Stufflebeam & Corey, 2014) consolida os objetivos e perspectivas de avaliação sob duas categorias básicas para o propósito da avaliação: as avaliações somativas e as avaliações formativas.

Uma avaliação somativa consiste em um processo *ex post facto*, ou seja, realizado após a conclusão de um programa – ou sua estabilização após um primeiro ciclo quando é um programa permanente – que estabelece retrospectivamente um juízo de valor acerca de seus resultados para os atores de interesse, em geral os próprios avaliados, os agentes do programa ou os seus patrocinadores e tomadores de decisão. Para Mark et al. (2000), a avaliação somativa corresponde ao objetivo principal de determinação do mérito – a qualidade do desempenho – e o valor – o alcance do efeito almejado – de um programa, devendo abranger as consequências reais de sua realização tanto em termos de resultados desejados quanto indesejados. Tal avaliação tem um caráter de prestação de contas para decisões acerca da manutenção, ampliação, correção ou encerramento dos programas (Scriven, 1991; Stufflebeam & Coryn, 2014).

Uma avaliação formativa contrasta com a somativa na medida em que é realizada de forma contínua e concomitante à implementação do programa com o objetivo de promover o seu aperfeiçoamento. Dessa forma, tem como principais interessados os agentes de

implementação, de maneira que se preocupa antes em fornecer informações tempestivas para correção dos rumos do programa do que em medidas de mérito e valor (Mark et al., 2000; Scriven, 1991). A avaliação formativa pode se dar sobre quaisquer etapas do programa, desde o planejamento até os resultados. No entanto, para que a avaliação dos resultados tenha valor formativo, esses devem ser antecipados e previstos, em geral, por avaliações intermediárias que permitam a compreensão das causas subjacentes oriundas do processo de execução, ou seja, quais etapas do processo são determinantes e devem ser monitoradas e eventualmente modificadas para garantia dos resultados – similar à perspectiva de processos proposta por Suchman (1967). Por isso, a distinção entre uma avaliação somativa ou formativa advém do propósito e tempestividade. Nas palavras de um pesquisador, citado por Scriven (1991, p. 169): “quando o cozinheiro prova a sopa, é formativo, quando o cliente prova a sopa, é somativo”. Por fim, segundo Mark et al. (2000), a avaliação formativa também pode apontar aperfeiçoamentos do próprio sistema organizacional mais amplo, com relação à sua capacidade de formular, desenhar, gerenciar e avaliar o próprio programa. Essa avaliação de desenvolvimento (*developmental evaluation*) visa modificar as próprias condições macro ambientais que interferem no programa (i.e. “trocar o cozinheiro ou o dono do restaurante”).

Além dos propósitos fundamentais das avaliações formativas e somativas outros teóricos levantam também outros objetivos para a realização de uma avaliação. Mark et al. (2000), por exemplo, denotam a função de controle de conformidade que as avaliações podem ter, na medida em que verificam o quanto uma atividade está alinhada a normas e regulamentações, cumpre com seus requisitos legais e alcança seu público-alvo de maneira justa e equânime. O autor alerta, no entanto, que essas avaliações não são orientadas para o mérito e valor, haja vista que um programa pode ser completamente regular, mas não ser efetivo.

Um segundo objetivo proposto por Stufflebeam e Coryn (2014) é que as avaliações servem para a disseminação de conhecimento e conscientização social da qualidade de serviços e produtos ou melhores práticas organizacionais, por exemplo, por meio da avaliação e ranqueamento de instituições educacionais. Por fim, ambos os autores apontam para o valor científico das avaliações, na medida em que permitem testar teorias e hipóteses no âmbito da dinâmica de programas e atividades organizacionais. Nesse contexto de pesquisa avaliativa, Lagbein & Felbinger (2006) ainda distinguem os desenhos descritivos de avaliação, com intuito diagnóstico – como as avaliações de necessidades – e as avaliações

com desenho causal que permitem inferências sobre as hipóteses de relações entre os componentes e resultados dos programas.

Uma outra decorrência dos esforços de sistematização dos pioneiros do campo de Avaliação foi permitir o enquadramento dos modelos de avaliação sob a perspectiva da Teoria Geral de Sistemas (Stufflebeam & Coryn, 2004). Na medida em que pressupostos de inter-relação com o contexto, orientação para objetivos e um fluxo de encadeamento entre insumos, processos e produtos em diferentes níveis é concebido, todas as características de um sistema aberto podem ser atribuídas a um programa (Katz & Kahn, 1978). Essa concepção torna-se possível a partir do estabelecimento dos componentes-chave para a estruturação integrada de múltiplas perspectivas de avaliação, como visto em Suchman (1967).

Nesse ínterim, modelos sistêmicos de avaliação foram desenvolvidos visando abranger tais perspectivas. Um dos modelos mais difundidos foi o CIPP, proposto por Stufflebeam em 1960, criado originalmente para o campo educacional e posteriormente adaptado para uma vasta gama de atividades e campos de conhecimento (e.g. programas sociais, ciências da saúde, atividades organizacionais, engenharia civil e projetos militares) (Stufflebeam, 2003). O CIPP se fundamenta em uma abordagem de sistemas sociais e compreende as avaliações como um processo cíclico e dinâmico (Stufflebeam e Coryn, 2014). Segundo Stufflebeam (2003), o acrônimo CIPP – contexto, insumo, processo e produto – refere-se aos conceitos-chave abrangidos por quatro tipos de avaliação. Avaliações de Contexto, “C”, referem-se ao diagnóstico das necessidades, problemas e oportunidades dentro de um determinado ambiente para que se possam definir objetivos e metas que, posteriormente, servirão de referência para a avaliação do mérito e do valor das atividades realizadas e dos benefícios alcançados. Os insumos, “I”, envolvem a avaliação de estratégias alternativas, planos de trabalho e custos envolvidos na implementação de um programa, permitindo aos avaliadores, planejar adequadamente os esforços, dimensionar propostas justificadas de financiamento, detalhar planos de ação e estabelecer critérios para priorização de alternativas de execução. As avaliações de processos “P” monitoram, documentam e avaliam as atividades – em perspectiva formativa – promovendo ajustes e melhorias, possibilitando o controle e prestação de contas dos programas. Por fim, o segundo “P” se refere a Produtos, que visa identificar e avaliar resultados de curto, médio e longo prazo, previstos e não previstos com a finalidade de subsidiar as decisões de continuidade, aperfeiçoamento ou interrupção de um programa (Stufflebeam, 2003).

Na medida em que os paradigmas sistêmicos foram incorporados em variados campos científicos (Rousseau et al., 2018; Scott, 1998), os modelos sistêmicos de avaliação de natureza genérica, como o CIPP, terminaram por subsidiar também modelos avaliativos nesses variados campos, sobretudo nas ciências sociais. Segundo Scriven (1991), a partir de uma teoria geral de Avaliação, a qual estipula relações entre categorias de componentes sistêmicos, os demais campos de conhecimento têm derivado teorias específicas de avaliação que encapsulam seus objetos, variáveis, teorias e hipóteses de pesquisa.

Dentre esses campos, o TD&E, na medida em que havia incorporado pressupostos da Teoria Geral de Sistemas a partir de Gagné na década de 1960, também passou a incorporar pressupostos dos modelos sistêmicos de avaliação, em especial o CIPP, proposto por Stufflebeam (Abbad et al., 2020; Borges-Andrade et al., 2012). Tendo em vista a definição ampla do que caracteriza um programa, o TD&E se enquadra nessa definição, uma vez que seus modelos podem ser aplicados a ações isoladas de Treinamento, mas também a um conjunto dessas atividades, concebido como um programa de TD&E, com objetivos e atividades bem definidas (Borges-Andrade et al., 2012).

Como visto na revisão do histórico de evolução do TD&E, muito embora o campo tenha estabelecido seus principais critérios de avaliação na década de 1950, a expansão do escopo de investigação a partir da década de 1960, envolvendo a análise da influência de características individuais, desenho instrucional e fatores contextuais, propiciou a incorporação das taxonomias – contexto, insumo, processo, produto – da literatura de Avaliação, para a especificação de modelos de avaliação em TD&E (Abbad et al., 2020). Dessa forma, ainda que haja uma grande quantidade de modelos de avaliação em TD&E na literatura (Passmore & Velez, 2015), suas estruturas e perspectivas de análise podem ser comparadas e analisadas à luz das teorias de base do campo de Avaliação. Com efeito, todos os pressupostos apresentados relativos ao encadeamento de resultados, relações com o contexto, propósitos formativos, somativos, administrativos e científicos da avaliação podem ser efetivamente observados na literatura de avaliação do campo de TD&E, o que permite caracterizar seus modelos efetivamente como uma teoria específica de Avaliação, conforme sugerido por Scriven (1991). Seu caráter específico será determinado pelo seu enfoque na aprendizagem, tendo em vista que os programas que implementa são de natureza educacional, em âmbito organizacional, com utilização de processos de tecnologia instrucional (Meneses et al., 2010; Pilati, 2006).

Como visto na seção relativa ao Sistema de TD&E, o Treinamento se inicia e se conclui por meio de processos avaliativos (Borges-Andrade, 2012). Assim como os modelos gerais de avaliação, parte da Avaliação de Necessidades que, para o caso do TD&E, se volta para o diagnóstico de necessidades, problemas e oportunidades (Stufflebeam, 2003) que possam ser solucionadas por meio da aprendizagem (Ferreira & Abbad, 2014). Havendo o planejamento e a execução dos treinamentos para solução dessas lacunas, uma série de avaliações visa, como descrito anteriormente, prover informações sobre os resultados das ações educacionais para indivíduos, equipes, organização e sociedade – ou seja, resultados imediatos, intermediários e finais – assim como sobre o próprio processo de TD&E. Acerca desse último conjunto de avaliações, a própria definição trazida por Abbad et al. (2020) sobre o processo de avaliação em TD&E reflete os propósitos gerais trazidos pelas teorias do campo da Avaliação em que:

A Avaliação em TD&E consiste em um processo que inclui coleta, mensuração, análise de dados e emissão de juízo e valor sobre a eficiência, eficácia e a efetividade de uma ação instrucional. Ou seja, consiste na mensuração e no julgamento de valor sobre determinado programa ou projeto instrucional/educacional (Abbad et al., 2020, p. 226).

Tendo por fundamento o paralelo com o campo de Avaliação, seus modelos, pressupostos, perspectivas e objetivos de análise, a próxima seção busca realizar uma análise dos principais modelos de avaliação em TD&E. O objetivo é verificar como se estruturam os modelos de avaliação, seus principais componentes, critérios e perspectivas de análise. Para tanto, a seção descreve 16 modelos de avaliação e delinea as suas perspectivas predominantes por meio de uma análise comparativa. Por fim, um recorte é realizado sobre como a eficiência tem sido teorizada e operacionalizada pelo campo de TD&E.

2.2.1 Modelos de avaliação em TD&E

Conforme descrito na trajetória de evolução do campo de TD&E, os modelos de avaliação são propostos de maneira mais sistemática a partir da década de 1950, com ênfase nos resultados do treinamento como forma de justificar os investimentos realizados na atividade (Bell et al., 2017; Kraiger & Ford, 2021). Consoante ao observado na literatura de Avaliação, o TD&E buscou estabelecer uma definição geral para os critérios de valor por meio dos quais os treinamentos deveriam ser avaliados.

Um dos primeiros modelos desenvolvidos para o campo, foi proposto por Kirkpatrick na década de 1950 e terminou por estabelecer as bases para uma taxonomia de objetivos –

critérios – para as ações de TD&E que perdura e fundamenta a literatura até a atualidade tendo se tornado um dos modelos mais disseminados no mundo corporativo (Abbad, Gama & Borges-Andrade, 2000; Alliger & Janak, 1989; Alliger et al., 1997; Kraiger, Ford & Salas, 1993; Passmore & Velez, 2015).

Kirkpatrick propôs quatro níveis fundamentais de resultados do treinamento, cada um consistindo em um tipo de avaliação. Os dois primeiros níveis referem-se aos resultados imediatos do treinamento, quais sejam: “Reação”, que contempla a avaliação da satisfação, percepção de utilidade e opiniões dos participantes acerca do treinamento logo após a sua realização; e “Aprendizagem”, que busca avaliar a aquisição por esses participantes dos conhecimentos, habilidades e atitudes (CHA) implicados no treinamento (Borges-Andrade et al., 2012). Os dois níveis finais correspondem a resultados de longo prazo, avaliados após certo decurso de tempo após o treinamento, sendo eles: “Comportamento no cargo”, que visa averiguar se o treinamento promoveu mudanças no comportamento dos indivíduos no trabalho, consistindo na transferência das competências aprendidas; e “Resultados” ou “Organização”, que, na medida em que o TD&E tenciona projetar o desempenho individual para apoiar o desempenho organizacional, visa analisar as mudanças que podem ter ocorrido na organização em termos de melhoria de processo, inovações ou outros benefícios de gestão de pessoas como redução de *turnover* ou melhoria do clima organizacional (Abbad et al., 2020; Borges-Andrade et al., 2012; Meneses et al., 2010; Tharenou et al., 2007). Posteriormente, Hamblin (1978) propôs um quinto nível de resultados que corresponde ao de “Valor Final”. O nível de valor final abrangeria alterações mais significativas nos resultados e desempenho agregado da organização, como o retorno social, retorno financeiro, além dos benefícios para o ambiente social e econômico mais amplo (Abbad et al., 2020; Borges-Andrade et al., 2012; Kaufman & Keller, 1994; Passmore & Velez, 2015). Com vistas à sumarização desses níveis fundamentais de avaliação em TD&E a Figura 4 apresenta as suas definições e ordenamento.

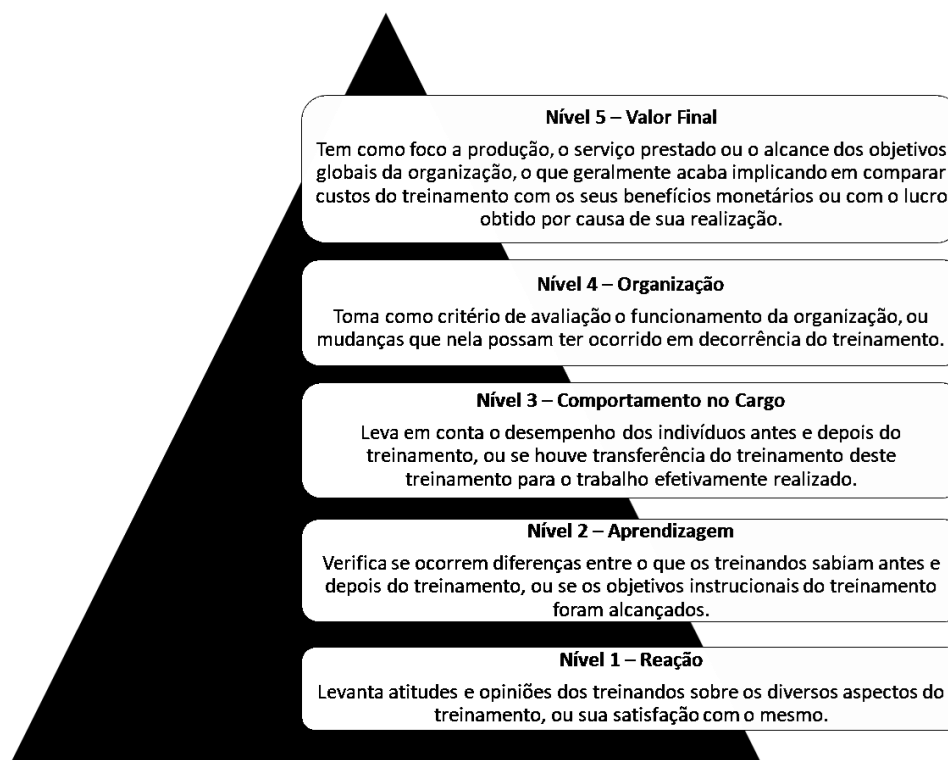


Figura 4. Níveis de Avaliação em TD&E

Fonte: Elaboração própria com base em Pilati & Borges-Andrade (2006).

Segundo Abbad et al. (2020), os modelos de Kirkpatrick (1976) e Hamblin (1978) são classificados na literatura de TD&E como modelos tradicionais e se caracterizam por focar os resultados das ações de treinamento. Os modelos tradicionais surgem com uma premissa de encadeamento de resultados, em perspectiva similar aos pressupostos da literatura de Avaliação. Dessa forma, assumiam que os níveis individuais de avaliação de aprendizagem e comportamento se constituíam em condições necessárias para a geração de resultados nos níveis da organização e valor final (Abbad et al., 2020; Holton, 1996; Meneses et al., 2010; Silva, 2011; Tharenou, 2010).

À medida em que o campo de TD&E passa a adotar uma perspectiva sistêmica com os marcos teóricos propostos por Gagné (1970), novos elementos de análise do treinamento são incorporados à literatura, tais como os insumos, características individuais e condições ambientais para sua efetividade (Borges Andrade et. al., 2012; Holton, 1996; Meneses et al., 2010; Meneses & Abbad, 2009; Passmore & Velez, 2015; Tharenou, 2010). Com isso, os modelos tradicionais tornaram-se insuficientes para a avaliação de TD&E, de forma que os modelos sistêmicos de avaliação de outras disciplinas, tais como o CIPP, fundamentam o desenvolvimento de modelos mais amplos de avaliação. Conforme Abbad et al. (2020), esses

são considerados os modelos integrativos, os quais se caracterizam por permitir a análise das variáveis que predizem ou interferem no alcance dos resultados do TD&E.

Ao longo do desenvolvimento do campo de TD&E, conforme pode-se observar em obras e manuais de treinamento contemporâneos, há uma grande variedade de modelos para avaliação de treinamento. Sob o ponto de vista científico, o paradigma sistêmico dos modelos integrativos prevaleceu na sistemática de avaliação, não obstante tenham, de forma geral, incorporado os níveis propostos por Kickpatrick e Hamblin que permaneceram como a taxonomia predominante de critérios de avaliação (Alliger et al., 1997; Kennedy et al., 2014; Passmore & Velez, 2015).

Além da compreensão dos modelos como tradicionais e integrativos, Borges-Andrade et. al. (2012) propõe a categorização dos modelos de avaliação em TD&E como genéricos ou específicos. Os modelos genéricos buscam delinear um *framework* que postula as relações entre os componentes e variáveis do TD&E, envolvendo seus insumos, processos, resultados e fatores contextuais, de forma a orientar as pesquisas no campo. Os modelos específicos, por sua vez, partem dos modelos genéricos, buscando investigar relações particulares entre determinadas variáveis e resultados.

Assim, o modelo CIPP, por exemplo, ao ter sido adaptado ao TD&E, afigurou-se como um dos modelos genéricos para o campo (Borges-Andrade et al, 2012). Já os modelos específicos se alinham ao que a literatura de Avaliação apontou como um dos seus possíveis propósitos, qual seja o de apoiar a investigação de hipóteses científicas específicas.

Por fim, cabe salientar que, à luz da literatura de Avaliação, também são referidos como somativos os modelos de avaliação em TD&E que aferem a eficiência, eficácia e efetividade dos treinamentos após a sua realização. Da mesma forma, ainda que menos comuns, o TD&E também prevê a realização de avaliações formativas que ocorrem durante o treinamento, sendo que nesse campo consistem, por exemplo, na testagem dos desenhos, estratégias e recursos instrucionais, realizando adaptações necessárias para a sua maior efetividade (Abbad et al., 2020).

Como será visto adiante, a despeito de sua ampla difusão, o modelo de base proposto por Kirkpatrick foi objeto de crítica e de complementação por diversos autores promovendo o surgimento de variados modelos de avaliação com maior ou menor grau de consonância com os critérios propostos pelos modelos tradicionais de avaliação (Holton, 1996; Kraiger, Ford & Salas, 1993; Passmore & Velez, 2015). Além disso, modelos e métodos de avaliação oriundos de outros campos, como o *Balanced Scorecard*, os modelos lógicos e a análise de

Retorno sobre o Investimento – ROI, vêm sendo adaptados ao campo de TD&E denotando a interdisciplinaridade do campo.

Assim, a análise proposta na presente seção tem como objetivo, primeiramente, descrever e caracterizar os principais modelos de avaliação contemporâneos utilizados em TD&E ressaltando os seus níveis e critérios de avaliação. A descrição dos modelos se baseia em revisão similar, proposta por Passmore & Velez (2015), de forma que foram selecionados 16 modelos de avaliação em TD&E, nacionais e internacionais, utilizados nas organizações e frequentemente mencionados na literatura. O delineamento não se pretende exaustivo, haja vista que novos modelos de avaliação são propostos continuamente por estudiosos do campo (Passmore & Velez, 2015).

Seguidamente à caracterização de cada um dos modelos, propõe-se utilizar-se dos pressupostos teóricos dos campos de Avaliação e de TD&E para delinear como os critérios e variáveis considerados em todos os modelos se associam, em uma perspectiva sistêmica, a componentes de insumos, processos, produtos, resultados e contexto. O objetivo da análise reside em compreender as principais dimensões, critérios e perspectivas utilizados para a análise da eficiência, eficácia e efetividade de ações e programas educacionais nas organizações. Como recorte proposto pela presente pesquisa, se analisará posteriormente como a perspectiva da eficiência é abrangida pelas avaliações realizadas no campo. Assim, a Tabela 1 apresenta uma compilação, com base em estudos teóricos, que relaciona e descreve os 16 modelos selecionados de avaliação de treinamento e seus critérios de avaliação conforme descrito na respectiva literatura. Cada um dos modelos é discutido e caracterizado para então serem analisados com relação às suas perspectivas de avaliação.

Tabela 1

Principais modelos de avaliação de treinamento do campo de TD&E

N	Modelo de Avaliação	Critérios de Avaliação	Aspectos Analisados
1	Modelos Tradicionais de Kirkpatrick (1976) e Hamblin (1978)	1 Reação	Opiniões em relação ao grau de utilidade e dificuldade do curso, bem como a satisfação com o treinamento.
		2 Aprendizagem	Aquisição dos conhecimentos, habilidades e atitudes previstos no treinamento.
		3 Comportamento	Mudanças no comportamento e no desempenho individual por meio da aplicação no trabalho dos CHA adquiridos nas ações de TD&E.
		4 Resultados	Mudanças na organização e no desempenho agregado tais como aumento de desempenho, redução de custos, melhorias de processo, redução de <i>turnover</i> e absenteísmo e melhoria do clima organizacional
		5 Valor Final (Hamblin, 1978)	Alterações no desempenho agregado da organização como o retorno social, retorno financeiro, além dos benefícios para o ambiente social e econômico mais amplo.
2	Modelo de Kaufman & Keller (1994)	1a Habilitação	Disponibilidade e qualidade dos insumos de recursos humanos, financeiros e físicos para o treinamento.
		1b Reação	Aceitação e eficiência dos métodos, meios e processos de ensino.
		2 Aquisição	Domínio e competência dos participantes.
		3 Aplicação	Utilização dos produtos do treinamento na organização.
		4 Resultados Organizacionais	Ganhos e contribuições à organização.
3	Modelo de Avaliação e Pesquisa de DRH (Holton, 1996)	1 Aprendizagem	Alcance dos resultados de aprendizagem almejados pela intervenção de desenvolvimento de pessoas, sendo influenciado por: reações, motivação e prontidão para aprender, atitudes no trabalho, personalidade e capacidade cognitiva. (Holton, 1996)
		2 Desempenho Individual	Mudança no desempenho individual como resultado da aprendizagem aplicada ao trabalho, sendo influenciada por: motivação para transferência, cumprimento de expectativas com o treinamento, desempenho no treinamento, atitudes no trabalho, condições (clima) de transferência e desenho do treinamento para transferência (<i>transfer design</i>) (Holton, 1996)
		3 Organização	Resultados no nível organizacional como resultado da mudança do desempenho individual, influenciado por: alinhamento aos objetivos organizacionais, expectativas de retorno financeiro e de valor, aspectos ambientais.
4	Modelo de Seis Estágios de Brinkerhoff (1998)	1 Definição de Objetivos	Avaliação das necessidades, problemas, fraquezas organizacionais e adequação e valor dos objetivos estabelecidos para o treinamento.
		2 Desenho do Programa	Avaliação da adequação do programa de treinamento, suas estratégias e materiais.
		3 Implementação do Programa	Monitoramento das atividades por meio das reações e satisfação dos participantes por meio de entrevistas e questionários.

		4 Resultados Imediatos	Avaliação do grau de aprendizagem dos participantes.
		5 Resultados Intermediários	Avaliação do grau de aplicação no trabalho dos conhecimentos obtidos.
		6 Impacto e Valor Final	Avaliação do valor do treinamento tendo em vista seus resultados e os custos e tempo despendidos,
5	Modelo CIPP aplicado ao Treinamento (Stufflebeam, 1978)	1 Contexto	Coleta de dados organizacionais
		2 Insumos	Planejamento e definição de alternativas para solução dos problemas, envolvendo orçamento, cronogramas, políticas e procedimentos.
		3 Processos	Implementação dos programas, reações, observação direta, entrevistas com participantes, grupos focais e análise de conteúdo.
		4 Produto	Grau de alcance dos objetivos de aprendizagem do programa de treinamento.
6	Modelo CIRO (Warr, Birdi & Rackham, 1970)	1 Conteúdos/Contextos	Avaliação do contexto organizacional e problemas operacionais de desempenho para proposição de treinamentos alinhados à cultura e clima organizacionais.
		2 Insumos	Alternativas de insumos e recursos mais adequados para diferentes tipos de intervenção.
		3 Reação	Experiência e sugestões dos participantes por meio de entrevistas.
		4 Resultados	Avaliação imediata (conhecimento, habilidades ou atitudes), avaliação intermediária (impacto, transferência e desempenho no trabalho) e avaliação final (impacto no desempenho departamental ou organizacional).
7	Modelo IPO (Bushnell, 1990)	1 Insumos	Indicadores de desempenho que influenciam a efetividade do treinamento, tais como qualificações dos participantes, desenho do programa, qualificação e qualidade de instrutores, materiais, instalações e equipamentos.
		2 Processos	Etapas de planejamento, desenvolvimento e execução do treinamento.
		3 Resultados/Produtos	Resultados imediatos (reações e desempenho de participantes) e de longo prazo (resultados organizacionais)
8	Modelo MAIS (Borges-Andrade, 1982)	1 Ambiente: Necessidades	Lacunas de CHA e desempenho individual e organizacional.
		2 Insumos	Fatores físicos e sociais e aos estados comportamentais e cognitivos, tais como características dos participantes (nível de escolaridade, idade, autoeficácia, cargo ocupado, comprometimento, estratégias de aprendizagem, crenças individuais, domínio prévio de competências para o treinamento, motivação para aprender e percepção de aplicabilidade do conteúdo aprendido).
		3 Procedimentos	Ações necessárias para facilitar ou produzir os resultados instrucionais ou a aprendizagem, tais como: sequência lógica e clareza dos objetivos instrucionais, adequação das aulas e materiais, uso de exercícios, discussões de casos e similaridade entre o treinamento e as situações de trabalho.
		4 Processo	Comportamentos dos participantes no treinamento envolvendo: desempenho em exercícios e testes, tempo dedicado, relações interpessoais entre aprendizes, seus tutores e coordenadores, motivação durante o treinamento, quantidade de desistentes e registros de atrasos e ausências.

		5 Ambiente: Apoio e	Características da organização e do local de vivência do indivíduo que afetam o treinamento, tais como parcerias, instalações e equipamentos para realização do treinamento, formação e experiência dos instrutores, expectativas em relação ao treinamento estímulo à transferência por parte do posto de trabalho, dos colegas e superiores.
		6 Ambiente: Disseminação	Meios e estratégias de comunicação de informações sobre treinamento, envolvendo: conteúdo das informações e o método, canais, fontes e natureza da divulgação.
		6 Resultados	Resultados imediatos do treinamento: reações e aprendizagem.
		7 Ambiente: Resultados em longo prazo	Mudanças esperadas e inesperadas, desejadas e indesejadas no comportamento individual, organizacional e medidas de valor final.
9	Modelo IMPACT (Abbad, 1999)	1 Suporte Organizacional	Opiniões relativas ao apoio gerencial ao treinamento, práticas organizacionais de gestão de desempenho e de valorização da força de trabalho.
		2 Características do Treinamento	Desempenho do instrutor, estratégias instrucionais e qualidade do material didático.
		3 Características da Clientela	Instituição, cargo, função, tempo de serviço, idade, nível de instrução, local de trabalho, quantidade de lotações de trabalho anteriores, interesse pessoal no treinamento e motivação para aprendizagem.
		4 Reação	Percepções sobre o treinamento envolvendo: qualidade das instalações, programação e material instrucional, desempenho do instrutor, assimilação e utilidade do conteúdo, expectativa de aplicação no trabalho, integração com outros profissionais e suporte organizacional.
		5 Aprendizagem	Assimilação dos conteúdos (Testes de conhecimento pré e pós-treinamento)
		6 Suporte à transferência	Condições para transferência envolvendo: reações favoráveis e apoio gerencial, da equipe de trabalho e da organização para transferência, disponibilidade de recursos e instalações adequadas.
		7 Impacto	Percepção de efeitos do treinamento no desempenho e motivação no trabalho.
10	Modelo IMTEE (Alvarez, Salas & Garofano, 2004)	1 Análise de Necessidades	Projeto e conteúdo do treinamento, mudanças nos treinandos e demandas organizacionais.
		2 Reação	Percepção de adequação, relevância e utilidade do treinamento e influenciado por características individuais dos treinandos.
		3 Transferência	Mudanças comportamentais no trabalho como resultado do treinamento, influenciado por: características do treinamento e da clientela
		4 Resultado	Mudanças quantificáveis em resultados organizacionais resultantes de mudanças de comportamento dos treinandos.

11	Modelos Lógicos Aplicado ao Treinamento (Kellog Foundation, 2004; McLaughlin & Jordan, 2015)	1 Recursos/Insumos	Recursos humanos, financeiros, organizacionais e comunitários disponíveis para implementação do programa de treinamento (Chyung, 2019).
		2 Atividades	Processos, ferramentas, eventos, tecnologias e ações que o programa implementa a partir de seus insumos e recursos visando resultados planejados.
		3 Produtos	Os resultados diretos das atividades dos programas.
		4 Resultados	Os conhecimentos, habilidades, atitudes, competências, status, nível de funcionamento modificados pelo programa e referentes ao nível individual.
		5 Impacto	Mudanças no nível organizacional, comunitário ou sistêmico proporcionados pelas atividades do programa.
12	Modelo ROI de Cinco Níveis de Philips (1991)	1 Reação e Satisfação	Reações dos participantes ao programa e a satisfação das partes interessadas com o programa e a implementação planejada. (Phillips, 2001)
		2 Aprendizagem	Mudanças em conhecimento, habilidades e atitudes relacionados ao programa de treinamento e sua implementação.
		3 Aplicação e Implementação	Mudanças de comportamento no trabalho e a aplicação e implementação específicas do programa de treinamento.
		4 Resultados de Negócio	Mudanças de impacto nos negócios relacionadas ao programa de treinamento.
		5 Retorno sobre Investimento	Comparação do valor monetário dos resultados de negócio com os custos do programa de treinamento.
13	Balanced Scorecard (Kaplan e Norton, 1996)	1 Aprendizado e Crescimento	Características da força de trabalho, por exemplo, satisfação, retenção, competências, nível de treinamento, conhecimentos-chave.
		2 Aprendizagem	Desempenho dos processos críticos para os resultados da organização, por exemplo, produtividade, qualidade, inovações, etc.
		3 Perspectiva do Cliente	Desempenho no mercado ou público-alvo da organização, por exemplo, satisfação, aquisição e retenção de clientes, fatia de mercado, etc.
		4 Perspectiva Financeira	Desempenho financeiro da organização, por exemplo, receita operacional, lucratividade, retorno sobre o capital, valor adicionado, etc. (Kaplan & Norton, 1996)
14	Métodos de Caso de Sucesso de Brinkerhoff (2003)	1 Foco e Planejamento da Avaliação	O que se deseja avaliar com relação ao treinamento.
		2 Criação do Modelo de Impacto	Definição dos resultados e benefícios potenciais almejados pelo treinamento.
		3 Taxas de sucesso	Identificação dos casos em que os objetivos foram alcançados.
		4 Entrevistas	Fatores de sucesso com base na opinião dos participantes para compreensão de como foram alcançados.
		5 Formulação de Conclusões	Recomendações com base na promoção dos fatores de sucesso, determinação do valor final e retorno do investimento do treinamento.

15	Avaliação de Escopo Total de Dessinger-Moseley (2010)	1 Avaliação Formativa	Juízos acerca do mérito e valor do processo de treinamento, das intervenções planejadas e implementadas e dos resultados imediatos e finais.
		2 Avaliação Somativa	Juízo acerca dos resultados imediatos como competências aprendidas e desempenho imediatamente após o treinamento.
		3 Avaliação Confirmativa	Resultados em longo prazo (pelo menos um ano após a intervenção) incluindo persistência da melhoria do desempenho, resultados e valor final para organização.
		4 Meta-Avaliação	Adequação, validade e confiabilidade dos instrumentos de avaliação utilizados em todo o processo.
16	Modelo SOAP-M (Passmore & Velez, 2012)	1 Autoavaliação (<i>self</i>)	Reações e auto percepção de comportamentos, personalidade e atitudes.
		2 Heteroavaliação (<i>other</i>)	Percepções sobre mudanças do comportamento e atitudes, e sobre a personalidade e desempenho do treinando.
		3 Conquistas	Resultados de desempenho individual e organizacional em metas e objetivos.
		4 Potencial	Mudanças no potencial individual por meio de desenvolvimento cognitivo e emocional.
		5 Meta-Avaliação	Comparação de resultados intergrupos com abordagem multimétodo.

Nota. Fonte: Elaboração própria com base em Borges-Andrade et. al., (2012); Brauschle & Schmidt (2004); Chyung (2015); Passmore & Velez (2015).

2.2.1.1 Modelos Tradicionais de Kirkpatrick (1976) e Hamblin (1978)

O primeiro modelo, pioneiro do campo de TD&E refere-se àquele proposto por Kirkpatrick na década de 1950 e que se tornou a principal referência para a avaliação de TD&E (Alliger et al. 1989) sendo logo em seguida complementado por Hamblin (1978). Tais modelos considerados tradicionais (Abbad et al., 2020) são o ponto de partida da presente revisão. Conforme discutido inicialmente, Kirkpatrick estabelece quatro níveis de avaliação de resultados, sendo complementado pelo quinto nível de Hamblin (1978), constituindo-se como critérios para avaliação: i) reação, que contempla opiniões acerca da utilidade, dificuldade e satisfação dos participantes com as ações educacionais, suas características e materiais instrucionais, desempenho de instrutores e condições de realização; ii) aprendizagem, que verifica de imediato o grau de apreensão de conhecimento, habilidades e atitudes (Kraiger, Ford & Salas, 1993) almejados pela ação educacional, normalmente com o intuito de verificar seu aproveitamento iii) comportamentos no cargo, acerca da utilização dos conhecimentos adquiridos pelos participantes em suas atividades de trabalho (transferência), possibilitando as mudanças de desempenho almejadas; iv) resultados, abrangendo efeitos agregados no nível organizacional, tais como aumento de desempenho, redução de custos, melhorias de processo, redução de *turnover* e absenteísmo e melhoria do clima organizacional; e v) valor final, relativo a efeitos do treinamento no desempenho final da organização ou benefícios mais amplos (Borges-Andrade et. al., 2012). Os primeiros dois níveis de resultados referem-se aos resultados imediatos dos treinamentos e representam o valor do treinamento para o indivíduo. O terceiro, quarto e quinto níveis de resultado são considerados efeitos em longo prazo, representando o valor do treinamento para a organização e para o ambiente econômico e social, avaliados após um determinado período desde o treinamento (Borges-Andrade, 2012; Passmore & Velez, 2015).

2.2.1.2 Modelo de Kaufman & Keller (1994)

O modelo de Kaufman & Keller (1994) originou-se a partir da crítica ao modelo de Kirkpatrick, acerca de sua simplicidade, a qual, segundo os autores, não permite a aplicação prática dos níveis 3 e 4 de desempenho, que integram a avaliação do treinamento ao contexto organizacional, limitando-se apenas aos níveis 1 e 2, relativos ao indivíduo (Passmore & Velez, 2015) – um aspecto confirmado empiricamente por outros autores (e.g. Kennedy et. al, 2014). Nesse sentido, de forma similar ao conceito cunhado por Hamblin (1978) de valor final, o modelo de Kaufman & Keller (1994) acrescenta o nível de resultados à sociedade e aos clientes

– chamado de nível mega-organizacional – de forma a levar em consideração os impactos no ambiente social, tanto positivos, como satisfação de clientes e reputação, quanto negativos, como poluição e riscos à sociedade (Kaufman & Keller, 1994; Passmore & Velez, 2015).

Conseqüentemente, implicando uma perspectiva integrativa para a avaliação, o modelo enfatiza a integração necessária com a etapa de Avaliação de Necessidades de Treinamento (ANT) e o planejamento prévio dos resultados do treinamento. Adicionalmente, os autores revisam os demais níveis, de forma que o primeiro nível de avaliação de reação deve também considerar a adequação e o valor percebido pelos participantes da qualidade e eficiência dos recursos utilizados no treinamento, entendendo que a satisfação está ligada à percepção positiva acerca desses fatores. Por fim, a avaliação de aprendizagem deve avaliar de imediato o comportamento, antecipando as chances de sua aplicação no trabalho para o alcance de resultados organizacionais. Nesse sentido, Kaufman & Keller (1994) ampliam o modelo de Kirkpatrick e estipulam com maior especificidade fatores de avaliação que abrangem mais claramente a ótica organizacional, desde a melhor utilização dos insumos até o resultado final para a sociedade.

2.2.1.3 Modelo de Avaliação e Pesquisa de DRH (Holton, 1996)

Holton (1996), baseando-se em uma crítica à ausência de estrutura e validade teórico-empírica do modelo de Kirkpatrick, propõe um modelo baseado em relações empíricas entre diferentes variáveis que levam a três níveis de resultados causalmente encadeados: aprendizagem, desempenho individual e mudanças na organização. A Figura 5 apresenta o modelo, seus fatores, as relações principais (setas em negrito) e secundárias que afetam resultados de treinamento.

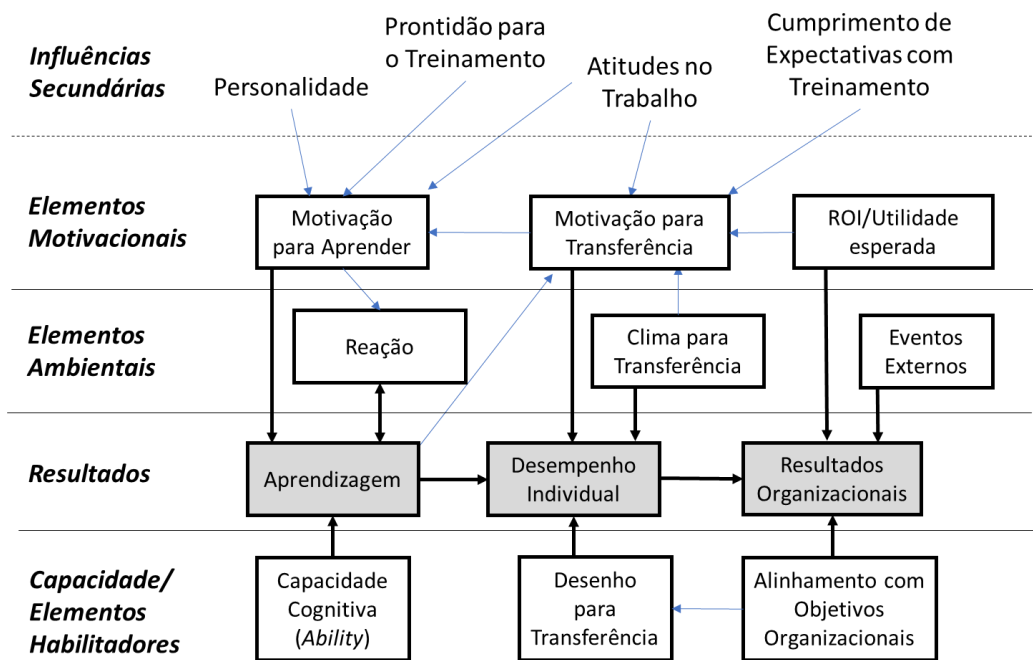


Figura 5. Modelo de Avaliação e Pesquisa em Desenvolvimento de Recursos Humanos
 Fonte: Holton (1996). Tradução própria.

Para Holton (1996) a avaliação deve ter uma abordagem científica, buscando validar modelos que considerem variáveis que expliquem os resultados do treinamento, sob pena de se incorrer em erros de avaliação caso se adote a abordagem mais simples de quatro estágios de Kirkpatrick. Uma primeira crítica reside na noção de que a reação como medida de satisfação (em suas palavras, uma medida de “felicidade”) não é um resultado do treinamento, mas um atributo do ambiente de aprendizagem (*learning environment*) que, juntamente com a motivação e a prontidão para aprender, propiciam o resultado, de fato, que é a aprendizagem. Essa por sua vez, não pode ser avaliada sem a consideração das características individuais dos aprendizes, tais como personalidade, atitudes no trabalho e capacidade cognitiva (Holton, 1996; Passmore & Velez, 2015).

Com relação ao desempenho individual, resultante do treinamento, deve-se avaliar as condições do indivíduo e do ambiente propícias à transferência – i.e. a aplicação do conhecimento adquirido no trabalho (Passmore & Velez, 2015) – considerando: a motivação para transferência, cumprimento de expectativas com o treinamento, desempenho no próprio treinamento, atitudes e as condições (clima) no trabalho para transferência e se o treinamento em si foi desenhado com vistas à transferência (*transfer design*) (Holton, 1996). Por fim, o impacto na organização deve levar em consideração, primeiramente, se o treinamento foi planejado com a análise do ambiente organizacional e dos seus objetivos, tendo em vista que

nem todos os treinamentos têm essa finalidade, de forma que, mesmo que apresente bons resultados individuais, isso não se refletirá na organização (Freitas & Mourão, 2012; Holton, 1996). Adicionalmente, o valor do investimento deve ser comprovado, de preferência antecipadamente, pela análise de retorno financeiro, mas também das expectativas organizacionais, que motivem o apoio da organização ao treinamento. Por fim, fatores externos – econômicos, sociais, institucionais – podem ter um impacto positivo ou negativo no treinamento e devem ser isolados para que se averigüe adequadamente o efeito do treinamento (Holton, 1996). Sendo assim, enquanto Kaufman & Keller (1994) estenderam o modelo de Kirkpatrick, Holton (1996) buscou preenche-lo, com uma grande quantidade de variáveis contextuais e intervenientes que visam avaliar se um treinamento de fato pode gerar aprendizagem, desempenho e por fim mudanças na organização (Passmore & Velez, 2015). Como mencionado, a abordagem proposta por Holton (1996) é consoante a um modelo específico, orientado para o desenvolvimento científico e testagem de hipóteses, e tece uma crítica similar àquela cunhada por Suchman (1967) de que as avaliações devem buscar rigor científico e integrar o máximo de elementos que se associem aos critérios de valor (Stufflebeam & Coryn, 2014), no caso, do treinamento.

2.2.1.4 Modelo de Seis Estágios de Brinkerhoff (1998)

Partindo também de uma tentativa de complementação dos modelos de avaliação em TD&E, o modelo de seis estágios de Brinkerhoff (1998) adiciona duas etapas anteriores aos níveis de Kirkpatrick, retroagindo os esforços de avaliação para o monitoramento dos processos do Sistema de TD&E, de forma a implementar uma perspectiva formativa para a avaliação (Passmore & Velez, 2015). Conforme Brinkerhoff (1998), a avaliação deve ter como primeiro nível de análise a etapa de ANT analisando a adequação de valor dos objetivos definidos para resolução dos problemas organizacionais. Como segunda etapa adicionada, Brinkerhoff (1998) propõe a avaliação do desenho do programa instrucional analisando sua adequação perante aos objetivos anteriormente estabelecidos, assim como a adequação de suas estratégias e materiais, aproximando-se da perspectiva de avaliação do mérito e do esforço (Mark et al., 2000; Suchman, 1967). Os níveis seguintes propostos por Brinkerhoff (1998) trazem uma subdivisão dos resultados individuais e uma compreensão relativamente distinta em relação aos modelos tradicionais. Para Brinkerhoff (1998), o terceiro nível de avaliação de Reação deve levantar a satisfação de participantes de treinamentos por meio de entrevistas e questionários, devendo ter, além de caráter somativo, caráter também formativo, como uma ferramenta de

monitoramento dos treinamentos. A partir daí o modelo adquire a perspectiva somativa similar aos modelos tradicionais, compreendendo três níveis de avaliação de resultados, os “Resultados Imediatos” que abrangem o nível de aprendizagem dos participantes do treinamento, os “Resultados Intermediários”, relativo ao desempenho no trabalho, e o de “Impacto e Valor Final” no qual se avalia o resultado e o valor do treinamento para a organização. Cabe destacar que no último nível Brinkerhoff (1998) insere uma perspectiva de eficiência para a avaliação, na medida em que postula que a análise deve considerar os resultados tendo em vista tempo e recursos utilizados (Passmore & Velez, 2015) que, de acordo com Dessinger e Moseley (2010) implica uma avaliação da sustentabilidade e retorno financeiro do treinamento.

2.2.1.5 Modelo CIPP aplicado ao Treinamento (Stufflebeam, 1978)

O CIPP, conforme anteriormente discutido, é proposto por Stufflebeam na década de 1960 com ênfase em contextos educacionais e difunde-se para outros campos, dentre eles o treinamento. A adaptação do CIPP para o campo de TD&E representa a sua transição para o paradigma sistêmico na avaliação (Abbad et al., 2020; Borges-Andrade et al., 2012) a qual também produziu outros modelos sistêmicos, como o CIRO, IPO e o MAIS, que serão vistos adiante. Uma vez que adotam explicitamente a visão sistêmica do treinamento, tais modelos se configuram como integrativos (Abbad et al., 2020) distinguindo-se pela consideração de fatores interferentes do ambiente. Além disso, ao envolver a avaliação integrada do programa de treinamento e de todo o seu ciclo, implicarão em uma perspectiva tanto somativa quanto formativa de avaliação, menos comum no campo de TD&E (Abbad et. al., 2020). Assim, os modelos CIPP, IPO e MAIS, por exemplo, operacionalizarão avaliações do próprio processo de treinamento e suas etapas, do planejamento à execução do curso (Abbad et. al., 2012, Passmore & Velez, 2015).

Segundo Chyung (2019), a adaptação do CIPP ao campo de TD&E implicou na consideração da ANT como parte dos esforços de avaliação do treinamento. Segundo a autora, essa inclusão é coerente na medida em que o TD&E é um processo cíclico, em que as avaliações de valor final se conectam com a ANT, retroalimentando o processo de treinamento. Nesse ínterim, a avaliação de Contexto e de Insumos do CIPP consistirá na etapa de pré-treinamento do processo de TD&E (Chyung, 2019), na qual, por meio da ANT realiza-se a análise do contexto organizacional, tarefas e indivíduos para definição de lacunas de aprendizagem (Salas et al., 2012). A etapa de insumos, mais especificamente, corresponderá à avaliação dos recursos e informações necessárias ao treinamento e de alternativas estratégicas para a resolução das

lacunas identificadas (Chyung, 2019). Os primeiros dois estágios possuem perspectiva formativa e como resultado das avaliações produzem políticas, orçamentos, cronogramas, propostas e procedimentos para implementação do plano de TD&E (Passmore & Velez, 2015).

O aspecto formativo e também distintivo do CIPP reside na avaliação no nível do “Processo”, em que se coletam informações para o monitoramento dos programas de treinamento visando correções tempestivas. Segundo Passmore & Velez (2015), as avaliações de reação integram, no CIPP, a etapa formativa de avaliação do processo, devendo ser realizada por meio de entrevistas, observação em sala de aula e grupos de foco para identificação de barreiras e dificuldades ao longo do treinamento. Por fim, a etapa de “Produto”, envolve a avaliação em perspectiva somativa, averiguando-se os resultados concretos do treinamento em termos de mérito e valor, ou seja, “examina o impacto do programa no público-alvo, bem como até que ponto o programa é sustentável e transferível” (Passmore & Velez, 2015, p. 142).

Segundo Passmore & Velez (2015), o desenho sistêmico do CIPP propondo a divisão do processo de Treinamento em etapas bem definidas e com questões específicas de avaliação facilita o processo de diagnóstico do próprio sistema de forma cíclica e contextual, o que permite melhorias no programa de treinamento, assim como em toda a organização. Dessa forma, o CIPP afigura-se como um modelo integrativo, formativo e somativo (Abbad et al., 2020; Stufflebeam & Coryn, 2014). Pode-se observar também que a consideração da ANT e da reação como avaliações de natureza formativa remetem, em muitos aspectos, ao que seria proposto também por Brinkerhoff (1998), evidenciando como tais pressupostos subsidiaram novos modelos no próprio campo de TD&E.

2.2.1.6 *Modelo CIRO (Warr, Birdi & Rackham, 1970)*

O modelo CIRO – Contexto, Insumo, Reações e Resultados (*Context, Input, Reactions and Outcomes*) – foi proposto por Warr, Birdi & Rackham (1970) e tem maior utilização em instituições de ensino, ou seja, no campo da Educação. No campo de TD&E, o modelo se caracteriza por realizar avaliações pré e pós treinamento para analisar suas margens de resultado e se destaca por prescrever abordagens qualitativas para melhor compreensão do contexto de aprendizagem (Passmore & Velez, 2015). A análise pré-treinamento, assim como o CIPP (Chyung, 2019), envolve uma avaliação do contexto organizacional – “C” – e dos problemas operacionais de desempenho, assim como a combinação mais adequada de insumos – “I” – entendidos como os recursos à disposição da organização, para definição de programas. Para a etapa pós-treinamento, o CIRO especifica expressamente a avaliação dos resultados em nível

individual por meio da avaliação de reação – “R” – enfocando, porém, a entrevista em profundidade e sugestões ao invés de questionários. Como quarto critério, delinea a dimensão de resultados (*outcomes*) – “O” – em três níveis: imediato (mudanças no conhecimento, habilidades ou atitudes, i.e. aprendizagem), intermediário (impacto, transferência e desempenho no trabalho) e final (impacto no desempenho departamental ou organizacional) (Passmore & Velez, 2015). Como pode ser visto, a perspectiva de avaliação do “Processo” de TD&E é pouco enfatizada nesse modelo, de forma que pode ser considerado um modelo integrativo e somativo.

2.2.1.7 Modelo IPO (Bushnell, 1990)

O modelo IPO proposto por Bushnell (1990) é mais recente que os demais modelos sistêmicos e foi desenvolvido no contexto das organizações privadas, voltado especificamente para o TD&E e não para o campo educacional, como inicialmente foram os demais modelos sistêmicos. O IPO se diferencia por não partir da análise do Contexto, mas dos Insumos seguindo para avaliação do Processo e posteriormente para as avaliações dos efeitos do treinamento em nível individual e organizacional (Figura 6).

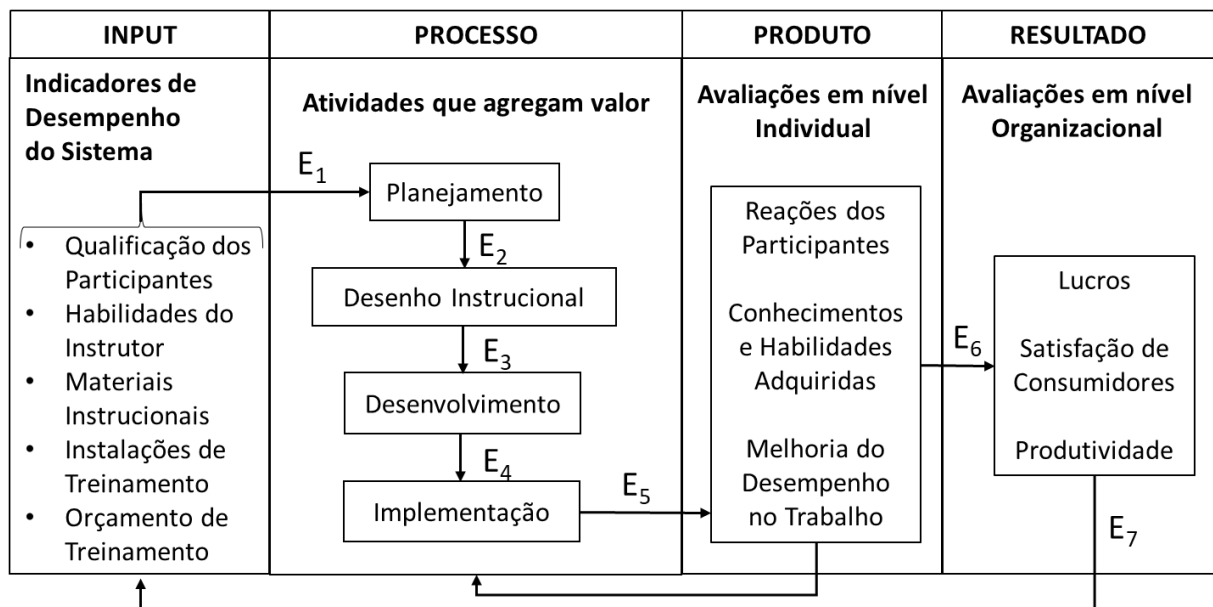


Figura 6. Uma abordagem insumo-processo-produto para a avaliação de treinamento.

Fonte: Bushnell (1990) com adaptações e tradução livre.

Bushnell (1990) define que avaliações intermediárias de natureza formativa devem ser realizadas ao longo desse processo (E₁ a E₇) visando identificar pontos de monitoramento para intervenção e melhoria. A etapa de avaliação dos Insumos (E₁) deve considerar todos os

chamados “Indicadores de Desempenho do Sistema”, entendidos como fatores com potencial para contribuir para a efetividade do treinamento. Tais fatores são propostos em pelo menos cinco categorias: qualificação dos participantes do treinamento, habilidades do instrutor, materiais instrucionais, instalações para o treinamento e o orçamento disponível para o treinamento, abrangendo, assim, recursos disponíveis, mas também elementos instrucionais (Passmore & Velez, 2015). A etapa de “Processo” implica a avaliação de como os treinamentos são planejados e implementados pela organização, devendo-se analisar se os pressupostos instrucionais, como a definição de objetivos educacionais adequados, estão sendo cumpridos. As avaliações em nível individual consistem na etapa de avaliação do “Produto” entendido como os níveis tradicionais do campo de TD&E de reação, aprendizagem e desempenho que informam recursivamente o processo instrucional. Por fim, as avaliações em nível organizacional abrangem aspectos como os resultados organizacionais em termos de lucratividade, satisfação de clientes e produtividade, retroalimentando a etapa de avaliação dos insumos. A premissa do IPO é monitorar cada etapa do processo de treinamento para, recursivamente, identificar precisamente onde podem estar ocorrendo problemas (Passmore & Velez, 2015).

2.2.1.8 Modelo MAIS (Borges-Andrade, 1982)

Na esfera nacional, em 1982, Borges-Andrade baseou-se em modelos genéricos sistêmicos, dentre os quais o CIPP, para propor o chamado modelo MAIS (Borges-Andrade, 2006a, Borges-Andrade et. al. 2012). Inicialmente visando o contexto educacional, passou ser aplicado em organizações buscando a avaliação detalhada dos componentes das ações e programas de treinamento, com especial ênfase nos seus condicionantes ambientais (Figura 7). O MAIS incorpora os benefícios atribuídos ao CIPP de avaliação do TD&E em etapas bem definidas com base na perspectiva sistêmica (Passmore & Velez, 2015). Ao considerar também o contexto, se diferencia, no entanto, por especificar as categorias de determinantes contextuais para o treinamento, em consonância com as preocupações levantadas, por exemplo, por Holton (1996).

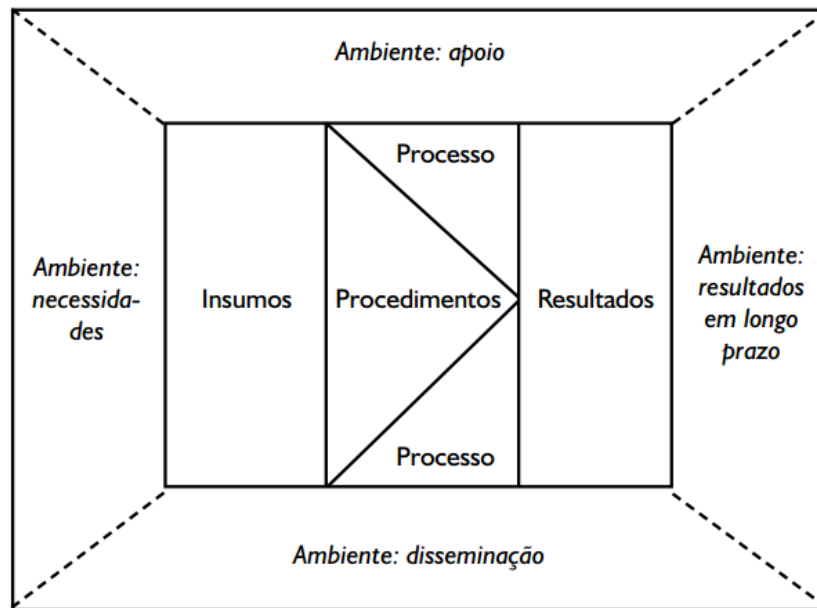


Figura 7. Modelo de avaliação integrado e somativo - MAIS

Fonte: Borges-Andrade (2006).

O MAIS incorpora a análise do contexto por meio do componente “Ambiente” de seu modelo, subdividido em quatro subcomponentes analisados antes, durante e após o treinamento: necessidades, apoio, disseminação e resultados de longo prazo (Borges-Andrade et. al., 2012). Sumariamente, o primeiro subcomponente concerne às “necessidades” resultantes da etapa de ANT e que justificam e dão origem ao treinamento, apresentando lacunas de aprendizagem e de desempenho individual e organizacional. O segundo subcomponente de “apoio” visa capturar a influência das condições fornecidas pela organização para a efetividade do treinamento, especificando-se fatores, tais como o fornecimento de instalações, equipamentos e instrutores com formação que propiciem a qualidade do treinamento, assim como a criação de ambientes de apoio à aprendizagem por parte de colegas, superiores, de desenhos adequados de trabalho e parcerias com outras organizações (Borges-Andrade et. al., 2012). O subcomponente “disseminação” envolve o estímulo ao treinamento, que segundo Borges-Andrade (2006), é sustentado por ações e estratégias de comunicação organizacional envolvendo a adequação e efetividade dos métodos, canais e conteúdos de divulgação de informações sobre treinamentos, os quais devem facilitar e motivar a busca da aprendizagem. Por fim, o subcomponente de “resultados de longo prazo” é considerado como parte da interface entre o TD&E e o contexto organizacional, e segundo Borges-Andrade et al. (2012) envolve os níveis finais de Kirckpatrick de desempenho individual e organizacional, ressalvado que considera efeitos desejados e indesejados como forma de aprimoramento do sistema de treinamento, similar ao proposto por Kaufman & Keller (1994).

Os “Insumos” no MAIS se referem sobretudo às características dos participantes – nesse aspecto aproximando-se de Bushnell (1990) e Stufflebeam (2003) – tais como: escolaridade, idade, autoeficácia, cargo, comprometimento, estratégias de aprendizagem, crenças, domínio de competências necessárias ao treinamento, motivação para aprender e percepção de aplicabilidade do conteúdo aprendido. Tais informações visam subsidiar o processo de aprendizagem e podem ter valor formativo, consoante pressupostos teóricos da psicologia instrucional (Borges-Andrade et al., 2012). O componente “Resultados”, por sua vez, avalia “o que é produzido pelos eventos ou programas de TD&E” (Borges-Andrade et al., 2012), compreendendo resultados imediatos de curto prazo, tais como a avaliação de reação e aprendizagem dos participantes do curso.

Com relação à avaliação do processo de ensino-aprendizagem, dois componentes distintos são definidos. O componente “Procedimentos” se detém sobre o próprio processo de planejamento instrucional, abrangendo a sequência lógica e clareza dos objetivos instrucionais, adequação das aulas, materiais e métodos de ensino e a similaridade entre o treinamento e as situações de trabalho (Borges-Andrade et al., 2012). Já o componente “Processos” visa realizar o monitoramento do comportamento dos aprendizes no contexto de treinamento e visa avaliar: o desempenho em exercícios e testes no treinamento; tempo dedicado; relações interpessoais e motivação durante o treinamento; quantidade de desistentes; e registros de atrasos e ausências.

O MAIS enquanto modelo genérico, serviu de base para uma série de estudos nacionais que desenvolveram escalas para mensuração de variáveis concernentes aos seus componentes. De acordo com Borges-Andrade et al. (2012, p. 21), o MAIS possui aspecto integrativo de relação com o ambiente, com enfoque somativo, mas também “apresenta rica contribuição para avaliações formativas, na medida em que prevê uma análise dos insumos, do ambiente e dos processos e procedimentos relativo às ações de TD&E”

Nota-se uma distinção no modelo MAIS quanto à classificação de alguns fatores em relação aos demais modelos sistêmicos em que, por exemplo, recursos organizacionais (e.g. instalações, qualificação de instrutores, investimento) e resultados de longo prazo são considerados como parte de um componente ambiental, do contexto organizacional, enquanto que em Bushnell (1990) – IPO – Warr et al. (1970) – CIRO – e Stufflebeam (2003) – CIRO – tais fatores compreendem respectivamente os insumos e os produtos ou resultados (*outcomes*) do treinamento. De forma geral, tais distinções podem se dever ao enfoque integrativo e instrucional que pode ser observado no MAIS em que um escopo maior de fatores interferentes acerca dos indivíduos e da organização é considerado em relação aos demais modelos sistêmicos, de forma a propiciar uma estruturação distinta do processo de avaliação.

2.2.1.9 Modelo IMPACT (Abbad, 1999)

Um dos modelos específicos desenvolvidos a partir do modelo genérico do MAIS, refere-se ao modelo IMPACT (Abbad, 1999), considerado um dos mais utilizados por pesquisadores e organizações nacionais (Borges-Andrade et. al., 2012). Como um modelo específico, o IMPACT foi desenvolvido com o intuito de testar relações empíricas entre variáveis que influenciam diretamente o impacto do treinamento com enfoque no nível do comportamento individual. O Modelo IMPACT afigura-se como um modelo integrativo e somativo ao partir dos pressupostos do modelo MAIS (Abbad, 1999). Porém, de forma distinta do modelo genérico de base, como um modelo específico, busca considerar as variáveis de contexto, insumo e processo em função da sua capacidade explicativa para os resultados enfocados da avaliação, no caso, o impacto do treinamento.

O modelo IMPACT constitui-se, portanto de sete componentes: suporte organizacional, características do treinamento, características da clientela, reações, aprendizagem, suporte à transferência existente na organização e, por fim, o impacto do treinamento no trabalho (Borges-Andrade et. al., 2012) (Figura 8).

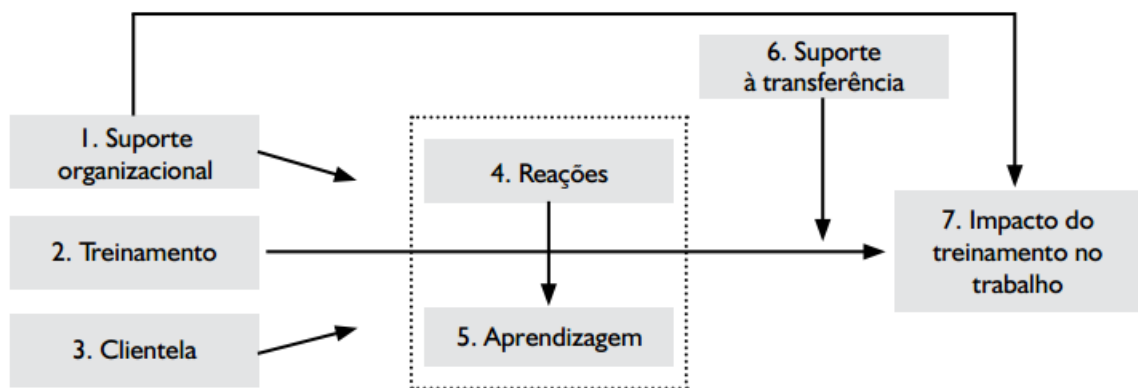


Figura 8. Modelo IMPACT de Avaliação do Treinamento

Fonte: Abbad (1999)

A avaliação do componente de suporte organizacional considera aspectos imediatos como o apoio gerencial para o treinamento, assim como a contribuição de políticas e práticas de gestão de pessoas, tais como a gestão de desempenho, relativa aos sistemas de metas, recompensas e orientação para o desenvolvimento profissional, e práticas de valorização da força de trabalho, como o respeito às ideias e interesses individuais (Abbad, 1999). O componente de características do treinamento delinea aspectos como natureza do objetivo do curso – classificado como instrumental (conteúdo geral) ou técnico/administrativo (conteúdo específico) – sua área de conhecimento, carga horária, qualidade do material didático e

desempenho e características do instrutor. Com relação à clientela, além dos dados pessoais e funcionais considera-se o interesse no treinamento e a motivação para aprender (Abbad, 1999).

O principal critério de impacto do treinamento no trabalho é avaliado em termos da percepção dos participantes acerca das mudanças decorrentes em seus desempenhos e motivação no trabalho. Com relação ao suporte à transferência esse é avaliado em termos do apoio de gestores e colegas e da disponibilidade de recursos e condições materiais para a transferência do treinamento (Abbad, 1999). Segundo Borges-Andrade et. al. (2012), dentre outras, a principal conclusão da testagem empírica do modelo é a relevância do componente de suporte organizacional ao treinamento, evidenciando uma dimensão até então pouco explorada de alinhamento entre treinamento e políticas, práticas e a cultura de gestão de pessoas na organização. O modelo também demonstrou como o suporte organizacional e as características do treinamento são preditores do impacto do treinamento e como o suporte e as características da clientela predizem os resultados para os níveis de reação e aprendizagem.

Dessa forma, o IMPACT, para além de sua capacidade de determinar o mérito e valor do treinamento, contribui também para os objetivos de uma avaliação que vise avançar o campo científico (Mark et al., 2000). Conforme há tempo foi apontado por Suchman (1967) e reiterado por Holton (1996), tal aspecto é relevante para o campo de TD&E, uma vez que modelos de avaliação pressupõem o encadeamento de objetivos de avaliação, devendo ser capazes de testar adequadamente essas relações para avanço da teoria específica sobre o fenômeno avaliado.

2.2.1.10 Modelo IMTEE (Alvarez, Salas & Garofano, 2004)

O modelo IMTEE, proposto por Alvarez, Salas e Garofano (2004), consiste também em um modelo específico de avaliação, conforme explicado por Borges-Andrade et. al. (2012). Construído a partir da análise da literatura de TD&E, o modelo se debruça sobre relevantes lacunas acerca da avaliação dos resultados em nível organizacional do treinamento. Segundo os autores, para o avanço do campo é preciso uma distinção entre a avaliação de treinamento e avaliação da efetividade do treinamento. A primeira estaria concentrada em resultados de aprendizagem e desempenho, em nível individual. A segunda se distingue por extrapolar o nível individual visando encontrar as relações do treinamento com o ambiente, sobretudo visando explicar quais fatores ambientais determinam o sucesso antes, durante e após o treinamento (Alvarez et al., 2004).

Os fatores contextuais para a efetividade do treinamento, consoante ao encontrado por Abbad (1999), incluem as características dos participantes, características do treinamento e características do ambiente organizacional. Para apreensão dos fatores ambientais, os autores

incluem a ANT como fator fundamental para a efetividade de treinamento e, seguindo a crítica ao modelo de Kirkpatrick feita por Alliger & Janak (1989) e Holton (1996), dentre outros autores, não consideram a reação como um resultado de treinamento, mas da ANT, ou seja, uma medida do quão apropriado ao contexto de trabalho foi o treinamento.

Conforme o modelo, os resultados propriamente ditos do treinamento são as mudanças de comportamento (nível 3 de Kirkpatrick), as quais decorrem de vários fatores: mudanças de atitude pós-treinamento como auto eficácia, desempenho no treinamento e aprendizagem cognitiva, influenciada pelas características individuais e do próprio treinamento (Alvarez, Salas & Garofano, 2004). Os comportamentos influenciam os resultados organizacionais (nível 4 de Kirkpatrick) mediados por aspectos e características organizacionais. As proposições do IMTEE se alinham a aspectos de modelos como o de Abbad (1999) e Holton (1996) ao enfatizar os efeitos intervenientes sobre a aprendizagem e a transferência, porém altera significativamente as relações entre os níveis de Kirkpatrick (Borges-Andrade et. al., 2012). Os autores concluem que avanços acerca da avaliação de efeitos em nível organizacional dependem da revisão e desenvolvimento de medidas mais apropriadas em TD&E (Alvarez et al., 2004).

2.2.1.11 Modelos Lógicos aplicados ao Treinamento (Kellog Foudation, 2004)

Conforme demonstrado anteriormente, modelos de outros campos foram incorporados à literatura de TD&E promovendo uma perspectiva ampliada de avaliação para além dos resultados abrangidos pelos modelos tradicionais. Parte desses modelos se originou na literatura de avaliação de programas, tais como o CIPP de Stufflebeam (2003), adaptados à avaliação de programas de treinamento. De forma similar, iniciativas recentes têm buscado a apropriação da abordagem de avaliação de Modelos Lógicos.

Os modelos lógicos foram desenvolvidos no contexto norte-americano pela Agência de Desenvolvimento Internacional dos Estados Unidos (*United States Agency for International Development – USAID*) e propostos inicialmente pela Kellog Foundation (2014) daquele país. Atualmente, a metodologia dos modelos lógicos é adotada principalmente por pesquisadores e profissionais para avaliação de programas públicos sociais e governamentais (Abbad et. al., 2012; Chung, 2019; Mourão & Meneses, 2012; McLaughlin & Jordan, 2015).

Segundo McLaughlin & Jordan (2015, p. 64) “um modelo lógico é um modelo plausível e sensato de como um programa funcionará sob certas condições ambientais para resolver problemas identificados”. Nesse sentido, o propósito de um modelo lógico, conforme sua proposição inicial pela Kellog Foundation (2014), era de que fosse utilizado como uma

ferramenta para o desenho de programas públicos para fins de planejamento e discussão dos recursos que deveriam ser mobilizados e das atividades que deveriam ser empreendidas para a resolução de um dado problema ou geração das mudanças e efeitos pretendidos.

A estrutura básica do modelo lógico apresentada pelo “Guia de Desenvolvimento de Modelos Lógicos” (Kellog Foundation, 2014) estabelecia uma relação entre o plano de trabalho e seu encadeamento em resultados almejados (Figura 9).

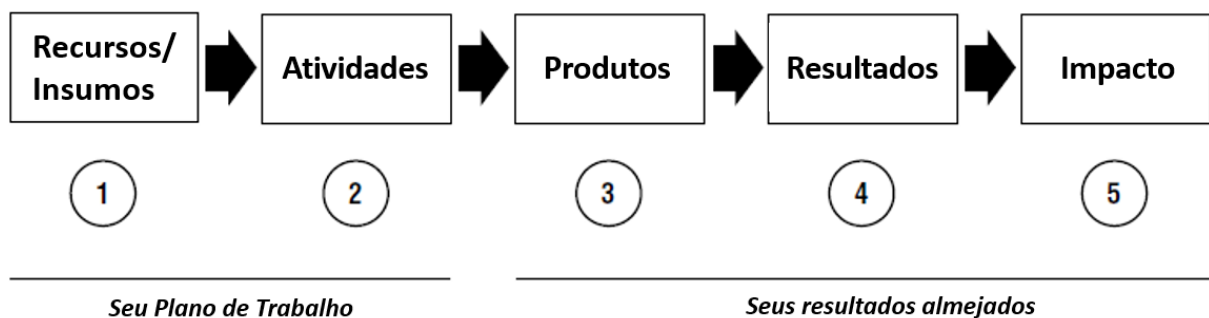


Figura 9. O Modelo Lógico básico
Fonte: Kellog Foundation (2004).

Os recursos, também chamados de insumos consistem nos recursos humanos, financeiros, organizacionais e comunitários disponíveis para a realização do programa pretendido. As atividades definem o que deve ser feito para alcance dos resultados almejados, delineados os processos, ferramentas, eventos, tecnologia e ações necessários à implementação do programa. Os produtos são as decorrências diretas das atividades podendo envolver tipos, níveis e metas de serviços a serem prestados pelo programa. Os resultados do programa são compreendidos em termos de mudanças específicas no comportamento, conhecimento, habilidades, *status* e desempenho de participantes do programa. Os resultados devem ser encadeados, definindo-se os resultados de curto prazo (1 a 3 anos) e como se relacionam a resultados de longo prazo (4 a 6 anos) e, por fim, ao componente final de impacto do programa (7 a 10 anos). Esse último componente corresponde à mudança fundamental intencional ou não intencional que ocorre como resultado final do programa no nível da organização, da comunidade ou do sistema. (Kellog Foundation, 2004).

O guia da Kellog Foundation (2004) especifica posteriormente como o modelo lógico deve ser usado para a gestão e avaliação dos programas. São propostas avaliações em perspectiva formativa para a gestão e somativa para atestar o alcance dos resultados. Nesse ínterim, considera fatores contextuais, riscos e alternativas de ação. Uma versão mais completa e atualizada do modelo lógico é apresentada por McLaughlin & Jordan (2015) visando cobrir, sobretudo, o contexto do programa (Figura 10)

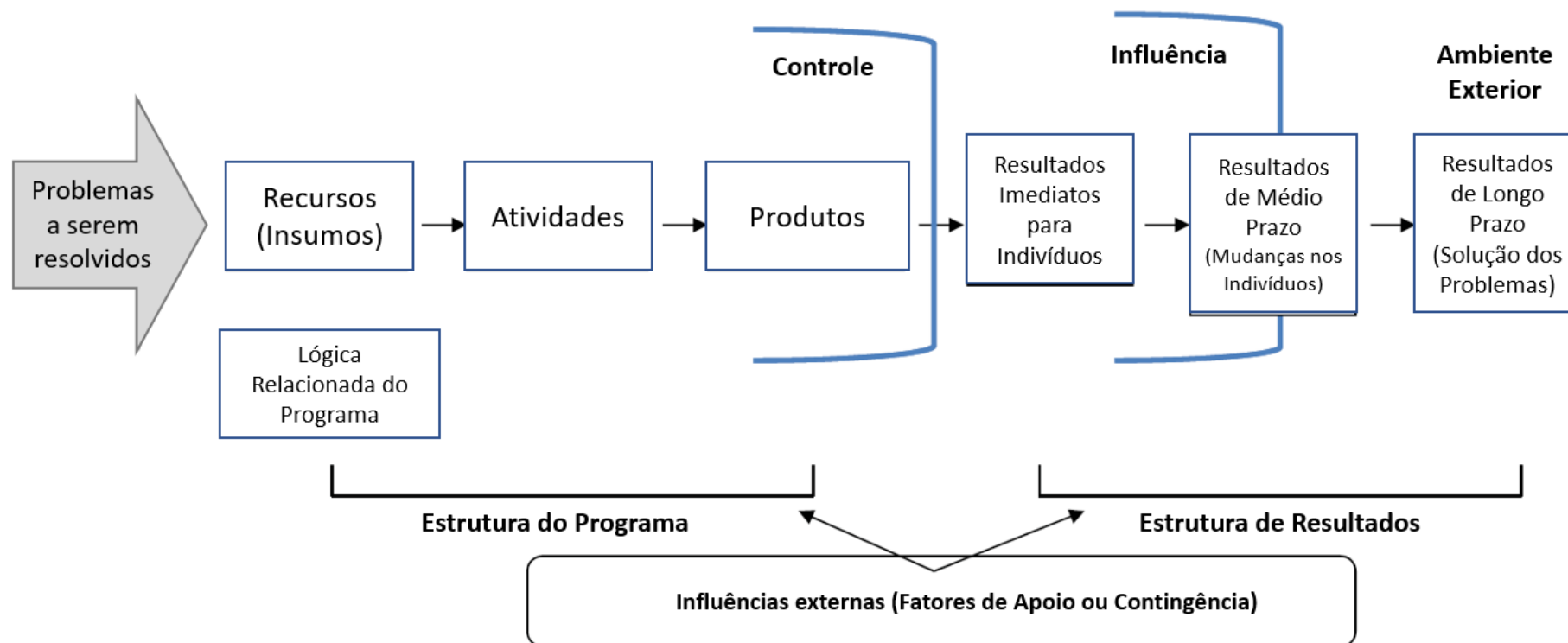


Figura 10. Estrutura do Modelo Lógico
 Fonte: McLaughlin & Jordan (2015).

Segundo Mourão e Meneses (2012) e Damasceno, Abbad & Meneses (2012), a aplicação dos modelos lógicos em TD&E deve partir de um conhecimento apurado do programa a ser avaliado para estabelecer um diagrama de relações causais entre variáveis de interesse com apoio de teorias científicas ou percepções de *stakeholders*. A diagramação dessas relações inicia-se pelo objetivo final, com a definição dos impactos almejados à organização ou à comunidade e então, retrospectivamente, são traçadas as relações entre variáveis, levantando-se quais resultados individuais levariam àquele objetivo, quais produtos e quais atividades de treinamento seriam necessárias. Por fim, definem-se os insumos que deverão ser operacionalizados para as atividades planejadas (Chyung, 2019; Mourão & Meneses, 2012).

A literatura nacional possui exemplos de aplicação dos modelos lógicos para a avaliação de treinamentos em diferentes contextos (Abbad et. al., 2012). Para ilustrar a diagramação hierárquica e sistêmica de um desses programas, a Figura 11 apresenta o modelo lógico desenvolvido por Meneses & Abbad (2009) com sete etapas para planejamento e avaliação de um curso sobre Desenvolvimento Regional Sustentável realizado por uma organização financeira acerca da concessão de crédito para aprimoramento de economias locais e elevação de índices de desenvolvimento humano. Como pode-se observar no diagrama lógico, utilizando-se como referências os componentes do modelo lógico, o seu impacto refere-se aos resultados organizacionais e sistêmicos na sociedade, com a definição de quais variáveis e indicadores devem ser analisados descritos pelas etapas 1, 2, 3 e 4 do diagrama (Cyung, 2019). As etapas representam respectivamente: (1) o impacto social final almejado pelo programa; (2) as mudanças organizacionais necessárias na organização financeira; (3) nas organizações produtivas locais; e (4) o aumento na integração de atores locais ao programa.

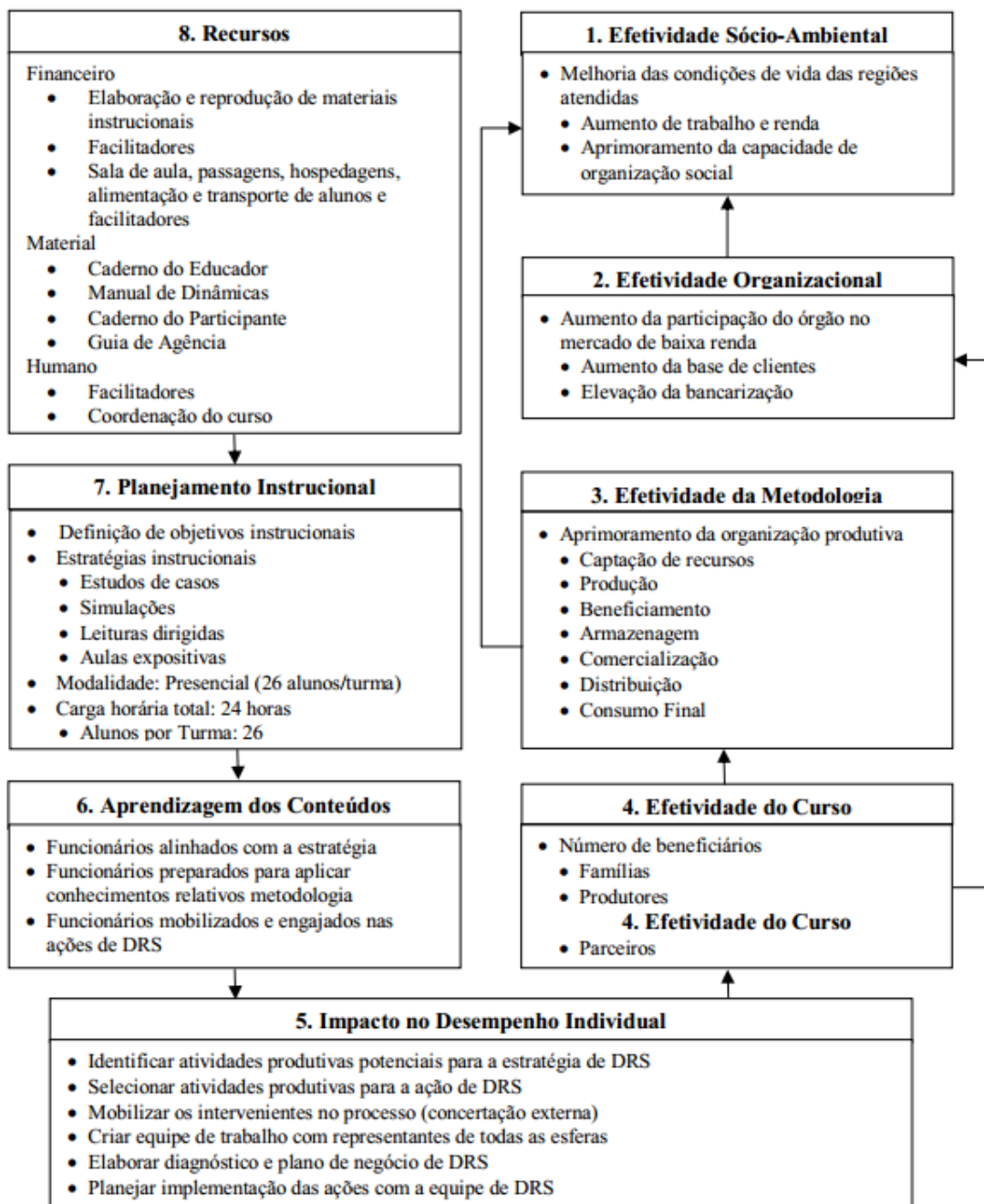


Figura 11. Exemplo de Modelo Lógico Final

Fonte: Meneses & Abbad (2009)

Seguindo-se os pressupostos para avaliação de programas públicos, o sequenciamento lógico das relações causais presumidas entre os componentes prevê que, para alcance das dimensões de efetividade, determinados resultados intermediários devem ser alcançados, referentes às mudanças de desempenho que os participantes do treinamento devem apresentar (5) que, por sua vez, dependem da adequada apreensão dos conteúdos ministrados (6). Tais elementos podem ser associados ao componente de produtos (*products*), relativo aos resultados

imediatos e diretos do programa e que representam também a perspectiva da eficácia do programa (Chyung, 2019).

Sob uma perspectiva de avaliação formativa a Etapa 7 aborda as ações do programa que visam desencadear os resultados mediatos e imediatos previstos no nível do indivíduo. Dessa forma, os componentes do programa são descritos, envolvendo os processos, ferramentas, eventos e tecnologias mobilizados a serem implementados como parte do processo de planejamento instrucional. Por fim, a etapa 8 define os recursos financeiros, materiais e humanos necessários ou disponíveis para as atividades, correspondendo aos indicadores de recursos e insumos utilizados para execução do programa (Chyung, 2019; Jannuzzi, 2005).

Os modelos lógicos permitem, portanto, o detalhamento de programas complexos, com múltiplos insumos, alternativas de execução e resultados (Abbad et. al., 2012). Em consonância com as teorias de base do campo de Avaliação (e.g. Suchman, 1967), os modelos lógicos possuem uma perspectiva sistêmica ao estabelecer o encadeamento de relações entre etapas e objetivos de um programa, além de definir variáveis contextuais que podem interferir na consecução do programa, favorecendo a sua avaliação (Meneses & Abbad, 2009; Mourão & Meneses, 2012; McLaughlin & Jordan, 2015). Sob o ponto de vista teórico Meneses & Abbad (2009), em atenção similar ao evocado por Alvarez et al. (2004), observam que os modelos lógicos facilitam a avaliação de treinamentos sob a perspectiva de sua efetividade nos níveis organizacional e social, ao especificar as variáveis e as relações que devem ser mensuradas e validadas, assim como os aspectos ambientais que devem ser considerados até o alcance dos objetivos nos níveis mais elevados.

2.2.1.12 Modelo ROI de Cinco Níveis de Philips (1991)

A busca pela efetividade no nível organizacional também se observa no modelo de retorno do investimento (ROI) de Philips que ganhou bastante relevância nas últimas décadas (Passmore & Velez, 2015). A premissa de Philips é de que, há uma motivação unânime em TD&E para a realização de avaliações de treinamento visando a otimização e justificativa dos investimentos nessa atividade e de que há uma crescente pressão nas organizações para que sejam apresentados resultados concretos para seus negócios, e não apenas para os indivíduos (Crochard & Davoine, 2011; Phillips, 1991). Assim, mantidos os quatro níveis de Kirkpatrick, Phillips (1991) propõe um quinto nível de valor final, relativo ao ROI (*Return on Investment*) do treinamento (Phillips, 1991). Este nível adicional consiste em métodos de análise de custo-benefício, para os quais Phillips (1991) especifica a necessidade de levantamento das chamadas “medidas duras” (*hard measures*) (Freitas & Mourão, 2012) relativas ao custo, tempo, materiais

e qualidade do processo produtivo e de provisão de TD&E. Como medidas perceptuais e coletivas (*soft measures*), o autor propõe a mensuração dos hábitos de trabalho (e.g. absenteísmo) clima organizacional (e.g. satisfação com o trabalho), sentimentos e atitudes, cultura de iniciativa, indicadores de desenvolvimento e avanço na carreira, e domínio de habilidades (Phillips, 1991). O cálculo do índice ROI mensura e converte os benefícios (resultados) do treinamento em valores monetários, de forma a compara-los aos custos totais do programa (Passmore & Velez, 2015). Concebido dessa forma, a metodologia ROI fundamenta-se em teorias econômicas que abrangem a dimensão da eficiência, na medida em que estabelece uma relação entre múltiplas variáveis de insumos e produtos sob o critério unificador do seu valor monetário (Phillips, 2010).

Segundo Passmore & Velez (2015) as principais fontes de críticas ao modelo recaem, primeiramente, sobre a necessidade de coleta extensiva de dados na organização (Brauchle & Schimdt, 2004; Passmore & Velez, 2015; Phillips & Phillips, 2010). Além disso, posto que a metodologia do ROI se baseia em uma razão entre investimento inicial e ganhos resultantes, o seu cálculo depende da tradução dos fatores em termos monetários (Phillips, 1991) o que nem sempre é possível em termos de resultados de projetos e programas (Mauss et. al., 2012). Por exemplo, vários autores apontam que os benefícios de treinamento são qualitativos, intangíveis, dificilmente mensuráveis ou rastreáveis como causa direta do treinamento (Freitas & Mourão, 2012; Passmore & Velez, 2015).

Dessa forma, Phillips (1991) estabelece que o ROI seja aplicado apenas a casos específicos onde haja a disponibilidade dos dados requeridos (Crochard & Davoine, 2011; Freitas & Mourão, 2012). Assim, o modelo ROI termina por ser recomendado para programas estratégicos de treinamento com claro foco em geração de resultado e de forma contextualizada, juntamente com medidas qualitativas de resultado – e.g. expectativas organizacionais – sendo fundamental que se busque isolar os efeitos do treinamento de outros fatores ambientais para seu cálculo adequado (Brachule & Schimdt, 2004; Holton, 1996; Phillips, 1991; 2010). Conclusivamente, o ROI abrange aspectos relacionados à eficiência, no entanto, possui limitações quando trata de variáveis não monetárias e quando o avaliador busca a análise de alternativas alocativas, detecção de gargalos e a influência de aspectos contextuais interferentes nos processos e resultados de um programa ou atividade organizacional (Buleca & Mura, 2014).

2.2.1.13 Modelo Balanced Scorecard (Kaplan e Norton, 1996)

O *Balanced Scorecard* (BSC), proposto por Kaplan & Norton (1996) também constitui um modelo que foi incorporado ao campo de TD&E. Desenvolvido no campo da Estratégia

Organizacional como um modelo generalista, tem ampla aplicação na avaliação da eficiência, eficácia e efetividade nas organizações (Buleca & Mura, 2014), e tem sido recepcionado como ferramenta para a avaliação da efetividade do treinamento (Baraldi & Cifalino, 2015. Brauchle & Schimdt, 2004; Silva & Abbad, 2011).

O BSC se baseia na definição de indicadores para monitoramento do desempenho estratégico de uma organização sendo que, originalmente estabelece quatro perspectivas de resultados encadeados: 1) aprendizado e crescimento 2) qualidade e eficiência de processos internos, 3) geração de valor ao consumidor e, por fim, 4) resultados financeiros, tais como o ROI e medidas similares (Kaplan & Norton, 1996).

O BSC tem recebido atenção no contexto acadêmico de TD&E como uma forma de operacionalizar avaliações de efetividade do treinamento, interligando as avaliações em nível individual e organizacional (Baraldi & Cifalino, 2015; Silva & Abbad, 2011). Em trabalho realizado por Silva & Abbad (2011), por exemplo, analisam-se os benefícios e limitações do BSC como método de avaliação do TD&E. Os autores confirmam a vocação da ferramenta para o alinhamento estratégico de aspectos relevantes para o campo de TD&E, assim como vantagens de voltar-se para além de medidas estritamente financeiras, tornando factível sua aplicação em organizações sem fins lucrativos. Porém, como também observado em outros estudos sobre eficiência nas organizações (e.g. Brauchle & Schmidt, 2004), Silva & Abbad (2011) ressaltam que as medidas e relações entre variáveis elencadas no BSC dependem das percepções dos tomadores de decisão. Dessa forma, podem ser relativamente arbitrárias e, caso mal especificadas e sem adequada testagem científica, as relações e medidas para os indicadores podem a levar a conclusões inadequadas acerca do desempenho (Silva & Abbad, 2011).

Um segundo estudo, de Baraldi & Cifalino (2015) utiliza o BSC, sob uma outra ótica de aplicação à gestão e avaliação de programas de treinamento. Segundo os autores, é possível o desenho de um *Training Scorecard* em que medidas relacionadas ao treinamento são definidas ao longo das quatro perspectivas do BSC, tais como a financeira (e.g. custo por hora de treinamento, gasto percentual da folha de pagamento em treinamento), do cliente (e.g. percentual e quantidade de treinados, satisfação com programas de treinamento), e dos processos internos (e.g. número de cursos, número de dias de treinamento, percentual de treinamento planejados realizados dentro do prazo). A partir dessa premissa os autores realizaram uma pesquisa-ação multimétodo junto a gestores e funcionários de 13 organizações privadas de pesquisa em saúde para levantamento dos principais indicadores e relações de causa e efeito que constituíam a efetividade de suas atividades de treinamento. A Figura 12 apresenta

os resultados da pesquisa, na qual quatro perspectivas de resultados foram definidas com base no modelo original BSC.

A perspectiva organizacional buscava definir como o treinamento deve contribuir para os resultados de negócio; a perspectiva dos aprendizes avalia – nos primeiros três níveis tradicionais de resultados – como o treinamento aperfeiçoa o desempenho individual; a perspectiva do processo interno averigua se a atividade de treinamento é realizada adequadamente; e, por fim, a perspectiva de aprendizagem e crescimento avalia se o sistema de treinamento recebe os recursos adequados para o seu funcionamento (Baraldi & Cifalino, 2015). Ao total são definidos mais de 40 indicadores distribuídos ao longo das quatro perspectivas e suas dimensões.

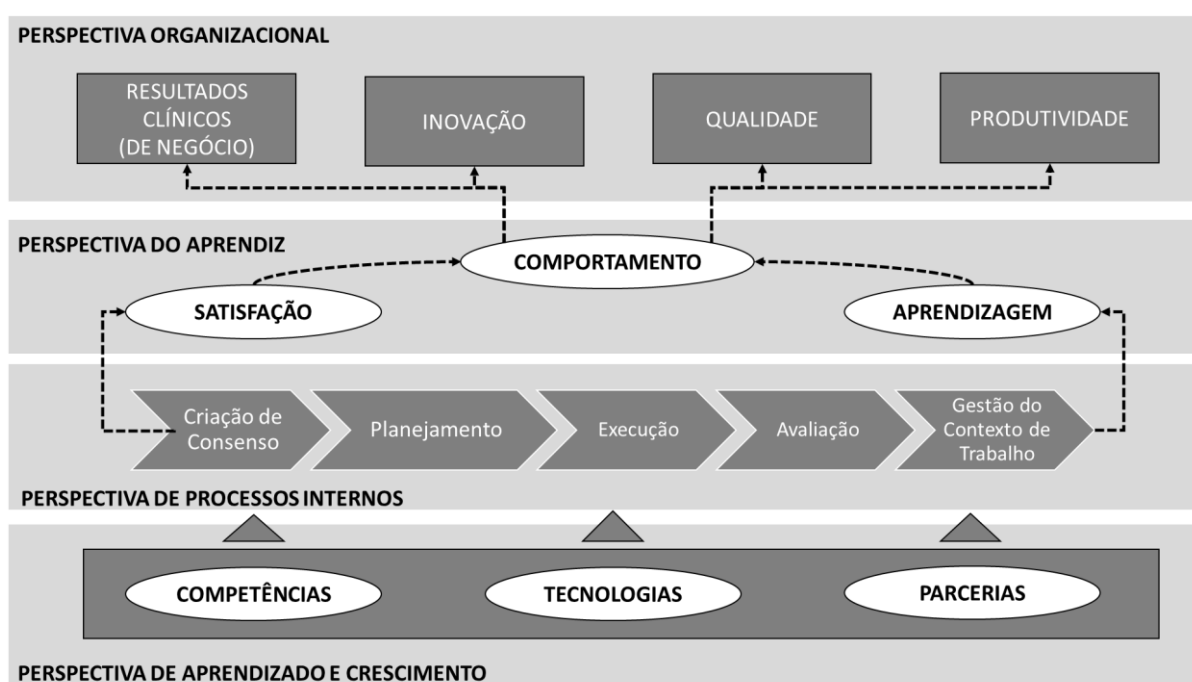


Figura 12. Mapa Estratégico do Treinamento

Fonte: Baraldi & Cifalino (2015). Tradução própria.

Para a primeira perspectiva, organizacional, os autores definem quatro dimensões abrangendo: resultados clínicos (intensidade média da dor do paciente e percentual de cirurgias com complicações), inovação (percentual de novos procedimentos implementados), qualidade (e.g. percentual de auditorias de saúde realizadas) e de produtividade (e.g. número de procedimentos de saúde por unidade de tempo). A segunda perspectiva voltada para os participantes de treinamento avalia as dimensões de satisfação (e.g. reação, utilidade, comparecimento), aprendizagem (e.g. pré e pós-teste de conhecimento) e comportamento (e.g. número de cirurgias realizadas).

A perspectiva de processos possui cinco dimensões correspondentes a etapas do processo de treinamento. Os autores incluem como etapa inicial, entendida como requisito para a efetividade de programas de treinamento, a construção de consenso e comprometimento entre gestores acerca da importância do desenvolvimento de competências, avaliando essa dimensão por meio, por exemplo, da quantidade de treinamentos previstos no planejamento estratégico da organização. Para as dimensões de planejamento, execução e avaliação variadas medidas buscam atestar a aplicação, qualidade, produtividade e custo de tecnologias de TD&E. Por fim, uma última dimensão de gestão do contexto de trabalho, visa detectar a existência de suporte à transferência por meio de um indicador de percentual de ações educacionais apoiadas (*follow-up*) (Salas et al., 2012). Por fim, a perspectiva de aprendizado e crescimento contempla indicadores a respeito das competências da equipe de treinamento (e.g. quantidade de horas de treinamento recebidas, conhecimento-chave e *turnover*), das tecnologias utilizadas (e.g. extensão do uso de ensino a distância e de simulações), e parcerias (e.g. percentual de ações educacionais realizadas em parceria com outras instituições).

As conclusões e vantagens apresentadas por Baraldi & Cifalino (2015) torna os benefícios comparáveis àqueles apontados também para os modelos lógicos, referindo-se, primeiramente à visualização diagramada de causas e efeitos em cascata e, em segundo lugar, a integração entre os níveis micro e macro organizacionais. Como diferenças, o BSC se restringe a indicadores-chave, não levando em consideração fatores contextuais e riscos de processo como fazem os modelos lógicos, sendo assim uma ferramenta de monitoramento, mas distanciando-se de um modelo de planejamento e contingência como os modelos lógicos. Por outro lado, o BSC avalia a atividade como um todo, não apenas programas particulares – os quais também podem ser avaliados por um BSC específico do programa (Baraldi & Cifalino, 2015) – e a insere, efetivamente, no contexto estratégico organizacional, contribuindo para uma perspectiva gerencial ampliada do TD&E, como também observado por Silva & Abbad (2011).

2.2.1.14 Métodos de Caso de Sucesso de Brinkerhoff (2003)

Seguindo-se à revisão dos modelos da literatura de TD&E, uma outra abordagem que se distingue metodologicamente das demais, a qual foi desenvolvida e bastante utilizada para departamentos de treinamento organizacionais, é o método de casos de sucesso de Brinkerhoff (Chyung, 2019; Passmore & Velez, 2015). Brinkerhoff assevera que 80% das falhas dos treinamentos advém de fatores não relacionados ao próprio treinamento, mas às características da organização e de seus participantes (Chyung, 2019) – algo que vai ao encontro de outros autores (e.g. Abbad, 1999; Holton, 1996; Phillips & Phillips, 2010). Visando identificar esses

fatores contextuais, por meio de abordagens qualitativas e quantitativas, o modelo de Brinkerhoff não busca mensurar impactos do treinamento, mas compreender, na opinião dos participantes, por que determinados treinamentos falham e outros são bem-sucedidos, quais estratégias foram utilizadas e quais condições estavam presentes (Chyung, 2019). O objetivo da metodologia é documentar, portanto, os casos de sucesso e relacionar os seus principais fatores determinantes para que sejam melhorados ou mitigados pela organização. Após essa etapa, em perspectiva quantitativa, os resultados tradicionais do treinamento são avaliados para que se tenha referência dos ganhos oriundos das estratégias de sucesso adotadas (Chyung, 2015; Passmore & Velez, 2015).

2.2.1.15 Avaliação de Escopo Total de Dessinger-Moseley (2010)

Também voltado para uma abordagem contextual e investigativa com aspectos qualitativos, o modelo de Avaliação de Escopo Total de Dessinger-Moseley (2010) (*Full-Scope Evaluation*) é considerado um dos mais recentes e se baseia em um fluxo interativo de análise do treinamento de maneira formativa, somativa e acrescenta aspectos chamados de confirmativo e meta-avaliativos (Passmore & Velez, 2015). A primeira etapa formativa avalia o mérito e o valor do processo de treinamento e de seus resultados como um todo. A etapa somativa contempla os resultados imediatos, como a aprendizagem ou desempenho, logo após o treinamento. Esses são observados novamente na etapa chamada confirmativa, realizada pelo menos um ano depois, para avaliar a persistência do desempenho e seu impacto e valor final para a organização. Essa etapa é multimetodológica, podendo abranger desde entrevistas até análises de ROI, correspondendo ao nível de resultados de longo prazo (Borges-Andrade et. al., 2012) similar ao nível final de Kirkpatrick ou mesmo níveis ulteriores propostos por modelos como os de Hamblin (1978) Kaufman & Keller (1994) e Phillips (1991). A inovação desse modelo reside na etapa de meta-avaliação que visa avaliar o mérito e o valor do próprio processo avaliativo, documentando-o extensamente, de forma que pode ocorrer simultaneamente às avaliações ou ao final do seu ciclo, enfocando nesse caso a sua validade e confiabilidade metodológica. A premissa é de que a avaliação não é uma atividade pontual, mas um processo contínuo devendo buscar a sua formalização na estrutura organizacional (Passmore & Velez, 2015).

2.2.1.16 Modelo SOAP-M (Passmore & Velez, 2012)

Por fim, o mais recente modelo citado por Passmore & Velez (2015) corresponde ao SOAP-M de sua autoria (i.e. Passmore & Velez, 2012). Partindo da crítica ao modelo de Kirkpatrick e dos

modelos desenvolvidos posteriormente – sobretudo por sua alta complexidade – o SOAP-M busca fornecer uma alternativa mais simples, prática e direta para sua utilização por departamentos de recursos humanos. O primeiro nível (*self*) consiste na autoavaliação de participantes de cursos imediatamente após sua realização, via entrevistas ou questionários, visando avaliar suas reações sobre o curso (Passmore & Velez, 2012; 2015). A etapa seguinte sugere a avaliação pelo supervisor, colegas e demais *stakeholders* do comportamento, da personalidade e das atitudes do treinando, podendo envolver testes pré e pós treinamento para avaliar mudanças e impactos de treinamento, assim como um contraste com os resultados da autoavaliação, oferecendo uma salvaguarda aos problemas metodológicos relacionados ao autorrelato, os quais, com efeito, são apontados em revisões do campo (e.g. Cassiano & Borges-Andrade, 2017; Garavan et. al., 2019). O terceiro nível corresponde aos níveis finais de Kirkpatrick e engloba resultados individuais e organizacionais sobretudo sobre metas objetivas de desempenho que devem ser avaliadas periodicamente. Aos quatro níveis tradicionais, acrescenta-se a avaliação de potencial, que deve utilizar instrumentos especializados pré- e pós-treinamento, para avaliar aspectos cognitivos e emocionais que evidenciem mudanças no potencial de crescimento individual (Passmore & Velez, 2015). Por fim, a meta-análise sugere a comparação sistemática, por múltiplos métodos, dos resultados de treinamentos entre grupos distintos de indivíduos, visando isolar efeitos contextuais para avaliação do treinamento (Passmore & Velez, 2012).

O modelo SOAP-M tem grande semelhança com as avaliações 360°, que são um método relativamente bem difundido de coleta informações de desempenho no qual indivíduos avaliam a si mesmos e são avaliados por seus pares, superiores e outros *stakeholders*. As vantagens dos métodos 360° destacadas por Brauchle & Schmidt (2004) referem-se ao alinhamento de expectativas de desempenho entre os indivíduos, fortalecendo equipes e aplicando-se sobretudo à avaliação de desempenho gerencial, dependente de aspectos comportamentais. O modelo SOAP-M visa, portanto, incorporar o espírito dessas práticas à avaliação de treinamento, adaptando os níveis de resultado tradicionais e acrescentando as medidas de potencial e valor final.

Sendo assim, a presente revisão dos principais modelos de avaliação do treinamento evidencia a riqueza teórica e empírica do campo de TD&E, assim como o progresso no desenvolvimento de modelos que partiram da proposta inicial de Kirkpatrick e buscaram estendê-lo com a adição de níveis de avaliação antecedentes e subsequentes, e aprofundá-lo por meio da consideração de variáveis intervenientes, tais como: características dos indivíduos e dos treinamentos, do processo de planejamento instrucional, do suporte organizacional à aprendizagem, suporte à transferência e das condições ambientais para a emergência de resultados individuais e organizacionais (Abbad et al., 2020).

Também é notável a interdisciplinaridade que o campo vem adquirindo com a adaptação de variados modelos e ferramentas de avaliação de outros campos de conhecimento tais como a gestão pública, economia, finanças e estratégia, como é o caso dos modelos lógicos, do *Balanced Scorecard* e dos modelos de impacto financeiro, assim como modelos destinados originalmente ao setor educacional e que foram transpostos para o TD&E como boa parte dos modelos sistêmicos, dentre eles o CIPP, o CIRO e o MAIS.

Como pode ser observado, os modelos tradicionais com enfoque em resultados nos níveis de reação, aprendizagem, comportamento, organização e valor final estabeleceram a taxonomia fundamental de critérios e objetivos de avaliação em torno da qual a grande maioria dos modelos de avaliação em treinamento se estruturou (Abbad et al., 2020). Não obstante, observa-se por meio dos modelos sistêmicos que perspectivas de avaliação de contexto, insumos, processos também se desenvolveram no campo. Buscando avaliar em que medidas tais perspectivas são abrangidas pelos modelos revisados, a próxima seção se deterá sobre como os aspectos fundamentais da eficiência, eficácia e efetividade são abordados por modelos de avaliação em TD&E.

2.2.2 Perspectivas de avaliação em TD&E

A avaliação consiste em um processo complexo e estruturado de análise que vai além da mensuração de indicadores, pois visa estabelecer um juízo de valor acerca dos resultados alcançados pela atividade realizada (Lagbein & Felbinger, 2006; Scriven, 1991). Segundo Abbad et al. (2020) os critérios fundamentais para as avaliações de treinamento se referem à sua eficiência, eficácia e efetividade. Para fins da análise de como essas perspectivas são abrangidas pelo campo, é preciso partir de uma definição distintiva tendo em vista que são termos frequentemente confundidos (Rosano-Peña, 2008).

Primeiramente, conforme Rosano-Peña (2008), a eficácia é uma medida normativa de fixação e alcance dos objetivos, sem se levar em conta os recursos empregados. Uma medida de eficácia compara um resultado obtido com a meta previamente estabelecida (Ferreira & Gomes, 2020; Mariano, 2007). A eficácia pode ser avaliada qualitativamente, em comparação aos objetivos previamente estabelecidos, ou quantitativamente por meio de uma razão entre os quantitativos alcançados e os quantitativos previstos (Mariano, 2007). Tomando-se o exemplo de encadeamento de objetivos de avaliação (Figura 3) proposto por Suchman (1967) e demonstrado na seção 2.2, a eficácia poderia se referir ao número ou proporção de pessoas que receberam os materiais informativos distribuídos pelo governo em um programa de conscientização em saúde. A eficácia também pode ter um caráter probabilístico quando, por exemplo, refere-se ao

percentual de proteção de um medicamento contra determinada enfermidade. No campo de TD&E, a eficácia se relacionará ao alcance dos objetivos educacionais definidos por um adequado processo de desenho instrucional para os cursos ofertados, refletindo a capacitação do público-alvo nas competências necessárias ao trabalho (Abbad et al., 2020).

A efetividade, sobretudo no contexto educacional, diz respeito aos efeitos de um programa na modificação dos comportamentos dos indivíduos e conseqüentemente na geração de mudanças para o contexto (Abbad et al., 2020). Retomando-se o exemplo de Suchman (1967), o objetivo final de um programa governamental de educação em saúde corresponderia, em nível individual, à mudança nos hábitos de saúde dos indivíduos para que se promova, em nível social, a redução da mortalidade. Aplicado ao campo de TD&E, o principal critério de efetividade diz respeito à modificação do comportamento dos indivíduos para aplicação em seu trabalho dos CHA que foram aprendidos nas ações educacionais, de modo a promover benefícios, maior desempenho e alcance de resultados para a organização e, potencialmente, para o contexto social mais amplo (Abbad et al. 2020; Aguinis & Kraiger, 2009; Cassiano & Borges-Andrade, 2017).

O que distingue a efetividade em relação à eficácia é que ela concerne a aspectos não controlados diretamente pelos agentes do programa (Manzoor, 2014), ainda que sejam os efeitos que o programa deseje produzir. Segundo McLaughlin & Jordan (2015), a eficácia, quando avaliada, permite inferir se a intervenção atende às condições para promover o efeito desejado em condições ideais. Já a avaliação da efetividade visa determinar se os resultados foram de fato alcançados nas condições reais em que foram implementados, as quais podem não ser ideais. Por esse motivo, é crucial a consideração de variáveis contextuais que interfiram na capacidade do programa gerar os comportamentos desejados nos indivíduos, assim como no encadeamento desses resultados até o nível contextual. Esse é o objetivo das perspectivas integrativas de avaliação (Abbad et al., 2020) e dos objetivos de muitos dos modelos revisados de TD&E, qual seja o de compreender os fatores que afetam a efetividade do treinamento. Os principais modelos específicos tratados nessa revisão tais como IMPACT (Abbad et al., 1999), o IMTEE (Alvarez et al. 2004) e o modelo de pesquisa em DRH (Holton, 1996), voltam-se especificamente para esse aspecto.

Dessa forma, a última perspectiva de avaliação, da eficiência, terá um papel bastante distinto das perspectivas de efetividade e eficácia. Segundo Rosano-Peña (2008, p. 85), a eficiência “visa assegurar a otimização da utilização dos recursos e, portanto, relaciona-se com os meios e não com os fins”. Aplicado ao TD&E, o critério da eficiência visa assegurar conforme Abbad et al. (2020, p. 227) “o máximo de resultados com o mínimo de fontes humanas, recursos financeiros e materiais”. Desta forma, é imprescindível para a avaliação da eficiência que se

levem em consideração os recursos utilizados em um determinado programa de treinamento, assim como a definição dos fins perante os quais a eficiência será avaliada. Por esse motivo, segundo Suchman (1967) a eficiência destina-se a avaliar, em perspectiva somativa, se determinado resultado foi eficiente tendo em vista os investimentos e esforços realizados. Assim, conforme Scriven (1991) a eficiência na utilização dos recursos integra uma perspectiva complementar à eficácia e à efetividade para a avaliação do sucesso de programas organizacionais:

A eficiência vai além da eficácia, trazendo uma referência à quantidade de recursos envolvidos. Implica a ausência de desperdício para uma determinada produção; ela [eficiência] pode ser aumentada pelo incremento de produto para um dado insumo. Isso não garante que os resultados sejam úteis. Por esta razão, em muitos contextos, os planejadores convencionalmente exigem que uma intervenção social deve ser tanto eficaz quanto eficiente (Scriven, 1991, p. 129, tradução livre).

Por esse motivo, assim como nos demais campos de conhecimento como delineado por Scriven (1991) e Suchman (1967), um programa integrado de avaliações deve buscar abranger conjuntamente os aspectos de eficiência, eficácia e efetividade. De forma geral, a eficácia e a efetividade de uma atividade, em termos dos critérios de valor, sempre serão a principal dimensão para determinação de seu sucesso (Scriven, 1991; Mark et al., 2000). Uma atividade que não alcança seus objetivos é imediatamente falha e ineficiente. Por outro lado, uma atividade completamente ineficiente ou inadequada do ponto de vista gerencial, não será capaz de gerar quaisquer resultados, de forma que há uma interdependência entre desempenho e eficiência (Lagbein e Felbinger 2006; Mark et al., 2000).

Tendo em vista tais pressupostos, a revisão dos modelos de avaliação em TD&E demonstra que variáveis relacionadas aos recursos se encontram presentes em parte deles, sobretudo os de natureza sistêmica. Visando esclarecer como esse componente é tratado pelos modelos, buscou-se categorizar os diferentes critérios de avaliação de todos os modelos de TD&E revisados a partir da arquitetura sistêmica genérica do CIPP e dos Modelos Lógicos para se averiguar como a perspectiva da eficiência pode ser abrangida por programas de avaliação em TD&E.

A Tabela 2 reapresenta os modelos revisados classificando seus critérios de avaliação nos componentes de Contexto, Insumo, Processo, Produto, Resultados de curto prazo, Resultados de longo prazo e tratando de maneira distinta os fatores interferentes apontados por alguns modelos. A comparação ordena os modelos a partir dos tradicionais, seguindo-se para os modelos específicos, e finalizando-se com os modelos sistêmicos. A Tabela 2 tem como inspiração análises realizadas por Abbad et al. (2020) e Nascimento (2018).

Tabela 2

Crítérios de modelos de avaliação em TD&E em abordagem sistêmica

MODELO		CATEGORIAS E VARIÁVEIS						
CIPP Stufflebeam (2003)	CONTEXTO	INSUMOS	PROCESSO		PRODUTOS			CONTEXTO
MODELOS LÓGICOS (McLaughlin & Jordan, 2015)	CONTEXTO (PROBLEMAS)	ESTRUTURA DO PROGRAMA			ESTRUTURA DE RESULTADOS			CONTEXTO
		INSUMOS	ATIVIDADES	PRODUTOS	RESULTADOS DE CURTO PRAZO (IMEDIATOS)	RESULTADOS DE MÉDIO PRAZO	RESULTADOS DE LONGO PRAZO (IMPACTO)	FATORES CONTEXTUAIS E AMBIENTAIS
Modelos Tradicionais de Kirkpatrick (1976) e Hamblin (1978)					Reação Aprendizagem	Comportamento (Desempenho)	Resultados Organizacionais Valor Final	
Modelo SOAP-M (Passmore & Velez, 2012)					Reação (Auto e Heteroavaliação)	Impacto no Desempenho Mudanças Individuais Meta-Avaliação (Comparação de Resultados)		
Modelo de Kaufman & Keller (1994)		Disponibilidade e qualidade dos insumos de recursos humanos, financeiros e físicos			Reação Aprendizagem	Transferência	Resultados Organizacionais Resultados Sociais	
Modelo ROI de Cinco Níveis de Philips (1991)		Custos totais do Treinamento			Reação e Satisfação Aprendizagem	Aplicação e Implementação	Resultados de Negócio; Retorno do Investimento.	

MODELO		CATEGORIAS E VARIÁVEIS						
CIPP Stufflebeam (2003)	CONTEXTO	INSUMOS	PROCESSO		PRODUTOS			CONTEXTO
MODELOS LÓGICOS (McLaughlin & Jordan, 2015)	CONTEXTO (PROBLEMAS)	ESTRUTURA DO PROGRAMA			ESTRUTURA DE RESULTADOS			CONTEXTO
		INSUMOS	ATIVIDADES	PRODUTOS	RESULTADOS DE CURTO PRAZO (IMEDIATOS)	RESULTADOS DE MÉDIO PRAZO	RESULTADOS DE LONGO PRAZO (IMPACTO)	FATORES CONTEXTUAIS E AMBIENTAIS
Modelo IMPACT (Abbad, 1999)	ANI Percepção de Suporte Organizacional	Características da Clientela	Características do Treinamento		Reação Aprendizagem	Impacto		Suporte à Transferência de Treinamento
Modelo IMTEE Alvarez et al. (2004)	ANT	Características Individuais	Características do Treinamento Desempenho do Indivíduo do Treinamento		Reação Auto Eficácia	Transferência	Resultado	Características Organizacionais
Modelo de Avaliação e Pesquisa de DRH (Holton, 1996)	Alinhamento aos Objetivos Organizacionais	Características dos Aprendizes	Desenho para Transferência.	Reação Motivação para Transferência	Aprendizagem	Desempenho Individual	Resultados Organizacionais	Clima para Transferência Expectativas de Retorno Financeiro e Valor Interferências Externas
Modelo Lógico aplicado ao Treinamento (Chyung, 2019; Damasceno, Abbad & Meneses, 2012)	Problemas e Objetivos do Programa	Recursos: humanos, financeiros, organizacionais e comunitários	Planejamento Instrucional Características do Treinamento		Aprendizagem	Mudanças nos Indivíduos	Impacto no contexto	
BSC aplicado ao Treinamento (Baraldi & Cifalino, 2015)	Criação de Consenso	Aprendizado e Crescimento: Competências da Equipe Tecnologias Parcerias	Processos Internos: Qualidade Produtividade Custo do processo		Satisfação Aprendizagem Comportamento		Resultados Organizacionais: Resultados de Negócio Inovação	Gestão do Contexto: Suporte à transferência

MODELO		CATEGORIAS E VARIÁVEIS						
CIPP Stufflebeam (2003)	CONTEXTO	INSUMOS	PROCESSO		PRODUTOS			CONTEXTO
MODELOS LÓGICOS (McLaughlin & Jordan, 2015)	CONTEXTO (PROBLEMAS)	ESTRUTURA DO PROGRAMA			ESTRUTURA DE RESULTADOS			CONTEXTO
		INSUMOS	ATIVIDADES	PRODUTOS	RESULTADOS DE CURTO PRAZO (IMEDIATOS)	RESULTADOS DE MÉDIO PRAZO	RESULTADOS DE LONGO PRAZO (IMPACTO)	FATORES CONTEXTUAIS E AMBIENTAIS
							Qualidade Produtividade	
Métodos de Caso de Sucesso de Brinkerhoff (2003)	Definição do Objeto para Planejamento da Avaliação.	Planejamento de Resultados Almejados	Criação do Modelo de Impacto:		Taxas de sucesso: Identificação de Casos de Sucesso Levantamento de Opiniões e Fatores de sucesso		Valor final ROI do Treinamento	Identificação de Fatores Contextuais
Modelos de Seis Estágios de Brinkerhoff (1998)	ANT Definição de Objetivos	Planejamento e alternativas instrucionais	Implementação dos treinamentos programados.		Aprendizagem	Transferência	Valor do Treinamento (custos e tempo despendidos).	
Avaliação de Escopo Total de Dessinger-Moseley (2010)		Avaliação Formativa: Processo de treinamento e seus resultados Meta-Avaliação: Instrumentos de avaliação utilizados em todo o processo.			Avaliação Somativa: Resultados para o Indivíduo e Organização Avaliação Confirmativa: Persistência dos Resultados e Valor Final			
Modelo CIPP (Stufflebeam, 2003) aplicado ao Treinamento	ANT Definição de Objetivos	Qualidade do planejamento Orçamentação Cronogramas Políticas Procedimentos	Monitoramento e Observação da implementação do programa		Alcance de Objetivos do Treinamento Impacto aos Beneficiários e Efetividade			

MODELO		CATEGORIAS E VARIÁVEIS						
CIPP Stufflebeam (2003)	CONTEXTO	INSUMOS	PROCESSO		PRODUTOS			CONTEXTO
MODELOS LÓGICOS (McLaughlin & Jordan, 2015)	CONTEXTO (PROBLEMAS)	ESTRUTURA DO PROGRAMA			ESTRUTURA DE RESULTADOS			CONTEXTO
		INSUMOS	ATIVIDADES	PRODUTOS	RESULTADOS DE CURTO PRAZO (IMEDIATOS)	RESULTADOS DE MÉDIO PRAZO	RESULTADOS DE LONGO PRAZO (IMPACTO)	FATORES CONTEXTUAIS E AMBIENTAIS
Modelo CIRO (Warr, Birdi & Rackham, 1970)	ANT Conteúdos	Alternativas de insumos mais adequados			Reação Resultados Imediatos (Aprendizagem)	Resultados Intermediários (Impacto para Indivíduos) e Finais (Impacto na Organização)		
Modelo MAIS (Borges-Andrade, 1982; 2006)	Ambiente: Necessidades	Características comportamenta is e cognitivas dos indivíduos	Procedimentos: Planejamento e Execução do Treinamento. Processo: Comportamentos dos Participantes no Treinamento		Reação Aprendizagem	Impacto no trabalho e na organização		Ambiente: Suporte e Disseminação
Modelo IPO (Bushnell, 1990)		Qualificação dos Treinandos e do Instrutor Recursos Financeiros Material Instrucional Instalações	Planejamento, desenvolvimento e execução dos treinamentos programados		Reações Aprendizagem Desempenho de participantes		Resultados Organizacionais: Lucro Satisfação do Consumidor Produtividade	

Nota. Fonte: Elaboração própria.

A análise proposta se baseia na interpretação dos objetivos dos modelos com base na sua descrição pelas suas fontes originais e pelas visões de outros autores (e.g. Borges-Andrade et al., 2012; Passmore & Velez). Assim, a Tabela 2 visou ilustrar as perspectivas abrangidas por cada um dos modelos, evidenciando nas células brancas os critérios principais que são enfatizados por cada um dos modelos, quando este é o caso. Nas células com tom cinza claro estão critérios de avaliação que, conforme os objetivos do modelo, são considerados de forma a apoiar o critério principal, ou seja, na medida que apoiam a explicação da variável-critério principal. Por fim, as células vazias coloridas em tom de cinza escuro são dimensões de avaliação não abrangidas pelos objetivos do modelo em questão.

Os primeiros modelos podem ser considerados como tradicionais somativos, segundo definição de Abbad et al. (2020), uma vez que possuem seu enfoque em critérios de resultados. Naturalmente, os primeiros modelos tradicionais de Kirkpatrick (1976) e Hamblin (1978) abrangem e delinham o que seria, na perspectiva dos Modelos Lógicos, a estrutura de resultados do TD&E, consistindo nos níveis fundamentais de reação e aprendizagem como resultados imediatos, comportamento como resultado intermediário e os resultados organizacionais e de valor final como resultados de longo prazo. Não se pode afirmar que há uma ênfase em qualquer desses critérios, mas resta claro que o modelo não abrange aspectos contextuais, de insumo, processo e produto como já discutido. O mesmo pode ser dito do modelo SOAP-M de Passmore & Velez (2012) que, muito embora seja o mais recente, apresenta critérios tradicionais, ainda que preveja inovações metodológicas como a avaliação 360° e a meta-avaliação. Os modelos seguintes também enfatizam os resultados, porém, como se vê no Modelo de Kaufman & Keller (1994), há um acréscimo de um critério de “Habilitação” que insere uma dimensão de avaliação da eficiência na utilização dos recursos do treinamento, ainda que de maneira perceptual. O modelo ROI de Philips (1991) também acrescenta um critério de valor final que considera subsidiariamente a dimensão dos insumos, no caso, os custos totais do treinamento, o que também permite abranger aspectos relacionados à eficiência. Nesse sentido, esses autores avançam na consideração dessa perspectiva de análise, no entanto, a circunscreve como parte de outras avaliações – de reação e valor final, respectivamente – não constituindo uma avaliação específica, por exemplo, de alternativas de alocação, como proposto por Suchman (1967), Scriven (1991) e Lagbein e Felbinger (2006).

Os modelos seguintes como o IMPACT (Abbad, 1999), o IMTEE (Alvarez et al., 2004) e Avaliação e Pesquisa em DRH (1996) são modelos específicos e somativos, ou seja, têm como objetivo original a investigação enfocada em um determinado critério de resultado,

respectivamente, o impacto do treinamento no trabalho, transferência (Borges-Andrade et al., 2012) e, no caso de Holton (1996), os níveis de aprendizagem, comportamento e resultados organizacionais (Passmore & Velez, 2015). Como modelos sistêmicos, trarão uma série de variáveis referentes ao contexto, resultados em nível individual e organizacional, além de variáveis de insumos e processo, na medida em que essas apoiam as hipóteses relacionais investigadas. Nessa perspectiva, porém, enfatizam hipóteses fundamentadas na teoria instrucional, considerando principalmente as características da clientela na dimensão de insumos, além das características dos treinamentos e variáveis de suporte e clima para aprendizagem.

Os Modelos Lógicos e o BSC, ainda que bastante distintos metodologicamente, também possuem um enfoque declarado em uma variável-critério, no caso os resultados em nível organizacional e ambiental. Dessa forma a estruturação das demais variáveis do modelo se volta para a explicação dos resultados de valor final. Não obstante, ambos os modelos abrangem múltiplas variáveis nas demais dimensões e permitem a realização de uma avaliação somativa, mas também formativa. Porém, como são modelos genéricos, para tanto é preciso que as variáveis do programa sejam adequadamente especificadas (Abbad & Silva, 2011; Mourão & Meneses, 2012). Com relação especificamente à dimensão de insumos, verifica-se que o BSC busca considera-los na perspectiva da “Aprendizagem e Crescimento”, mas com enfoque em recursos humanos, não abrangendo, pelo menos teoricamente, outros recursos como os materiais e financeiros. Por outro lado, os modelos lógicos abrangem explicitamente os recursos humanos, materiais e financeiros na dimensão de insumos. Dessa forma, dentre os modelos até então descritos o Modelo Lógico é o primeiro a apresentar uma estruturação em que a avaliação da eficiência pode ser realizada como estudo de alocação ótima de recursos em TD&E.

Os modelos de avaliação de escopo total (Dessinger e Moseley, 2010) e de seis estágios (Brinkerhoff, 1998) são modelos que se destacam por possuir um claro aspecto formativo ao lado do aspecto somativo. Nesse sentido, não se pode afirmar que enfatizam determinada dimensão. Dessinger e Moseley (2010) em seu *handbook* enfatizam alguns aspectos contextuais relevantes, porém, pelo menos dentre os critérios de avaliação descritos (Passmore & Velez, 2015; Brinkerhoff, 1998), não especifica claramente se considera antecedentes e variáveis contextuais interferentes no treinamento, de forma que podem não ser modelos integrativos. Da mesma forma, não há a consideração de recursos, relativamente às variáveis de insumos, sendo esses referentes às informações necessárias ao processo instrucional.

O modelo de casos de sucesso de Brinkerhoff (2003), por outro lado, é claramente somativo e orientado para resultados, porém distingue-se dos demais por sua abordagem essencialmente qualitativa. Considerando-se a descrição realizada por Passmore & Velez (2015), devido à sua metodologia flexível, esse modelo poderia, em tese, considerar qualquer critério para avaliação, caso seja eventualmente mencionado como um fator de sucesso pelos entrevistados, o que permitiria abranger qualquer dimensão de avaliação. Porém, em princípio, essa não seria uma avaliação sistemática e, em última análise, o modelo considera como casos de sucesso os treinamentos que alcançaram seus objetivos, isto é, os resultados almejados.

Finalmente, os modelos CIPP, CIRO, IPO e MAIS são modelos genéricos fundamentados em aspectos sistêmicos similares. Por definição, como modelos genéricos, não se pode estabelecer que haja uma determinada dimensão enfatizada. Não obstante, o CIRO conforme mencionado, não postula a avaliação na perspectiva dos processos, enquanto que o IPO não prevê avaliações acerca do contexto. O CIPP e o MAIS, por sua vez, são os modelos mais abrangentes, sendo que o MAIS, apoiado na literatura empírica, especifica com maior clareza e em grande número os tipos de variáveis que devem ser consideradas em processos de avaliação de TD&E. Com relação à dimensão de insumos e a perspectiva de avaliação da eficiência, nota-se que todos os modelos preveem variáveis de insumo. O MAIS, como discutido, especifica os insumos em termos de contribuições para o processo instrucional. Os recursos humanos, financeiros e materiais são mais diretamente avaliados como Insumos pelos modelos CIPP, IPO e CIRO, sendo, ao lado dos Modelos Lógicos, os que podem fornecer subsídios mais evidentes para avaliação da eficiência na relação entre produtos e recursos utilizados na provisão de ações educacionais.

Conclusivamente, considerando-se as perspectivas abrangidas pelos modelos de avaliação em TD&E, verifica-se que, justamente em função do objetivo de se atestar o mérito e o valor dos efeitos de aprendizagem que um programa educacional pretende gerar, o campo de TD&E desenvolveu majoritariamente modelos de avaliação somativa, com ênfase na eficácia e efetividade dos treinamentos, mas sem um enfoque metodológico específico para a avaliação da eficiência, ainda que esta seja reconhecida como uma das perspectivas fundamentais de análise. Isso não implica que os insumos não sejam relevantes para os modelos de avaliação, mas apenas que, conforme Suchman (1967), naturalmente, os critérios de sucesso – mérito e valor – de um programa são seus resultados, não a forma como utiliza seus recursos, de forma que, como explica Scriven (1991), a efetividade deve sempre ser o principal foco dos esforços de avaliação.

A perspectiva da eficiência possui um papel distinto, porém complementar em um programa de avaliação, apoiando a justificativa dos investimentos realizados, a detecção de desperdícios, problemas operacionais e gerenciais e ineficiências alocativas, os quais podem, em última análise, impactar os resultados potenciais e a sustentabilidade de um programa (Suchman, 1967). A próxima seção, visa, portanto, aprofundar a discussão acerca dessa lacuna da literatura de avaliação em TD&E cotejando a sua operacionalização.

2.2.3 A perspectiva da eficiência na avaliação em TD&E

A literatura de avaliação em diferentes campos, de forma geral, define que, junto à eficácia e efetividade de um programa, também é relevante a avaliação de sua eficiência (Scriven, 1991; Mark et al., 2000; Lagbein & Felbinger, 2006; Rosano-Peña, 2008). No entanto, conforme demonstrado na análise de variados modelos de avaliação em TD&E, a eficiência não consiste em um critério específico de análise e apenas parte dos modelos considera variáveis de insumo e, dentre esses, poucos abrangem os recursos dispendidos com ações educacionais.

Na medida em que o processo de aprendizagem ocorre em um contexto mais amplo de um sistema organizacional (Borges-Andrade, 1996), distintos critérios de valor, para além da dimensão instrucional, podem suscitar novas formas e objetivos de avaliação de um programa de treinamento. Para o caso da eficiência, os treinamentos em âmbito organizacional possuem uma dimensão gerencial, em que parte das decisões versam sobre as formas de alocação dos recursos disponíveis em alternativas de ações instrucionais.

Como visto no presente referencial, uma das principais motivações citadas pela grande maioria dos estudos sobre avaliação de treinamento é a grande quantidade de investimentos realizados nessa atividade pelas organizações. Consequentemente, há grande relevância no desenvolvimento de medidas que permitam justificar e orientar as decisões sobre os recursos investidos na qualificação da força de trabalho (Passmore & Velez, 2015; Markoulli et al., 2017). A avaliação de eficiência, em última análise, permite às unidades de treinamento lidar com situações de escassez de recursos às quais são frequentemente submetidas (CIPD, 2019), evidenciando quais aspectos podem ser otimizados, sem a perda da qualidade dos resultados, uma vez que a avaliação de eficiência não serve aos propósitos de promover cortes, mas de otimizar processos existentes (Scriven, 1991; Suchman, 1967).

Salas et al. (2012), em um importante artigo sobre o campo de TD&E, estabelecem com clareza quais as principais atividades e condições encontradas pela literatura para a efetividade das ações de treinamento nas organizações. Com relação à eficiência, no entanto, não há uma

sistematização similar para o melhor uso dos recursos de TD&E. Dessa forma, nos moldes da literatura consolidada de avaliação da efetividade, um aprofundamento teórico acerca da eficiência poderia promover um maior conhecimento de utilidade científica e prática para a definição de quais atividades e condições propiciam o alcance do máximo de resultados do treinamento com a menor utilização possível de recursos.

Como discutido anteriormente, parte dos modelos de avaliação em TD&E permite uma avaliação com essa perspectiva orientada para os recursos. Uma concepção inicial, portanto, de como podem ser integrados esses modelos de natureza sistêmica para o delineamento de uma avaliação de eficiência é apresentada na Figura 13. A Figura busca alinhar o processo sistêmico de TD&E com os componentes de avaliação do CIPP e do IPO e define, com base nos achados na literatura dos modelos de TD&E, os principais fatores contextuais antecedentes e condicionantes da efetividade de treinamentos. As variáveis encontradas na revisão dos modelos de avaliação são dispostas em termos de um Modelo Lógico baseado em McLaughlin & Jordan (2015).

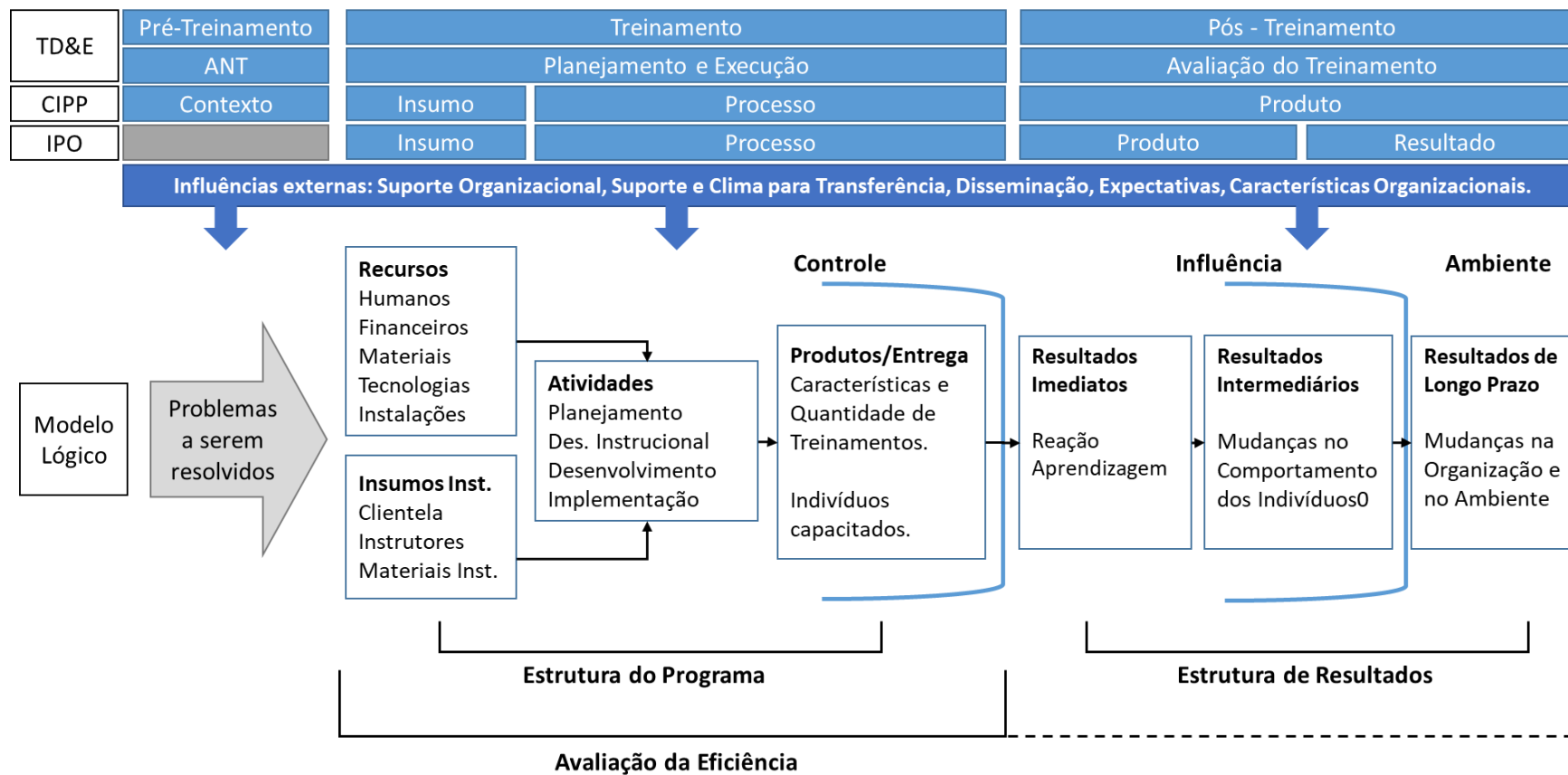


Figura 13. Integração de Modelos Sistêmicos para avaliação da eficiência em TD&E

Fonte: Elaboração própria com base em Borges-Andrade (2006), Salas et al. (2012), Stufflebeam (2003), Bushnell (1990) e McLaughlin & Jordan (2015).

O escopo fundamental para a avaliação da eficiência proposta é a estrutura do programa, uma vez que os processos e os produtos são realizados em função dos recursos disponibilizados para o treinamento. Uma distinção é realizada entre os “Recursos” propriamente ditos – humanos, financeiros, materiais, tecnologias e instalações – e os “Insumos Instrucionais” que são definidos para avaliações relacionadas à efetividade do treinamento, tais como as características da clientela, formação dos instrutores e qualidade dos materiais instrucionais. Em tese, tais aspectos também podem ser objeto de uma avaliação da eficiência, em que se averigüe se tais fatores podem influenciar também a eficiência de um treinamento – por exemplo, um público-alvo ou um instrutor que seja mais qualificado, que tenha determinado nível de poder econômico ou que resida em locais distantes podem ensejar treinamentos com maior ou menor necessidade de recursos para sua operacionalização ou reduzir a eficiência por comprometer o alcance dos resultados, implicando perda relativa de recursos. No entanto, no que tange à definição de alternativas de investimento, os indicadores de recursos são mais apropriados por operacionalizar diretamente essas variáveis.

Com relação ao processo, compreende-se que a avaliação da eficiência seria de natureza somativa, realizada após os treinamentos, de forma que deve analisar o processo de TD&E visando quantificar os esforços realizados, em lugar do que normalmente se observa na literatura em que esse componente é avaliado em perspectiva formativa, com relação à adequação do processo instrucional. Por fim, o componente Produtos permite avaliar se os treinamentos entregues condizem com os esforços e recursos definidos nas etapas anteriores. Não obstante, a avaliação da eficiência pode estender-se até a inclusão da estrutura de resultados para indivíduos, organização e sociedade (linha pontilhada), visando atestar os retornos do treinamento face aos recursos utilizados.

Cabe salientar que a literatura de TD&E tem buscado abordar a questão da eficiência sob diferentes perspectivas. Ford (1993), por exemplo, define métricas para o TD&E como forma de se realizar o *benchmarking* nessa atividade. Dentre essas, de forma similar às medidas propostas por Baraldí & Cifalinò (2015) na construção do BSC, duas medidas de eficiência do treinamento são propostas: custos do treinamento por estudante/hora; e custo da equipe de treinamento. Com relação à atividade de TD&E, propõe mensurar: investimento em TD&E por empregado, percentual da folha de pagamento gasta com TD&E; média de horas passadas em treinamento; percentual de empregados treinados por ano; e tamanho relativo da equipe de TD&E. Conforme Brauchle & Schmidt (2004), indicadores dessa natureza alinham-se a alguns métodos tradicionais de cálculo da eficiência, tais como os de custo-eficácia, custo-efetividade e custo-utilidade. Além desses, Brauchle & Schmidt (2004) elencam o BSC e métodos de

análise financeira, tais como de cálculo da taxa interna de retorno e o próprio ROI, o qual é incorporado ao TD&E por meio do modelo de avaliação de Philips (1991).

A crítica feita aos modelos e indicadores de eficiência dessa natureza é que esses não refletem a complexidade do sistema ou programa avaliado como um todo (Brauchle & Schmidt, 2004). Na medida em que se limitam a variáveis monetárias ou indicadores compostos de razões entre duas variáveis, deixam de considerar outros aspectos relevantes e, nesse último caso de Ford (1993), consistiriam antes em uma mensuração de resultados e não em um método de avaliação (Lagbein & Felbinger, 2006).

A relativa limitação acerca de como a eficiência é tratada no TD&E pode ser suprida com a incorporação de modelos de avaliação oriundos de campos como o da Economia (Lagbein & Felbinger, 2006), cujos modelos de análise subsidiam a avaliação de ações, programas e atividades em variados campos de conhecimento (Emrouznejad & Yang, 2018). O conceito de eficiência no âmbito da Economia vai além da relação insumo-produto trazida por indicadores operacionais unidimensionais (Rosano-Peña, 2008). A eficiência é tratada como um fenômeno complexo multivariado e multideterminado, que reflete uma propriedade latente dos sistemas produtivos relativa à sua capacidade de transformação de insumos em produtos e resultados (Bogetoft & Otto, 2011).

Segundo Bogetoft & Otto (2011), a avaliação da eficiência busca considerar o máximo de insumos, produtos e variáveis contextuais – chamadas de variáveis não-discricionárias, posto que estão fora do controle do gestor – envolvidas em um processo produtivo e a contribuição gerencial para o desempenho desse sistema. A Figura 14 apresenta um modelo básico de avaliação da eficiência, conforme proposto pelos autores sob uma perspectiva sistêmica.

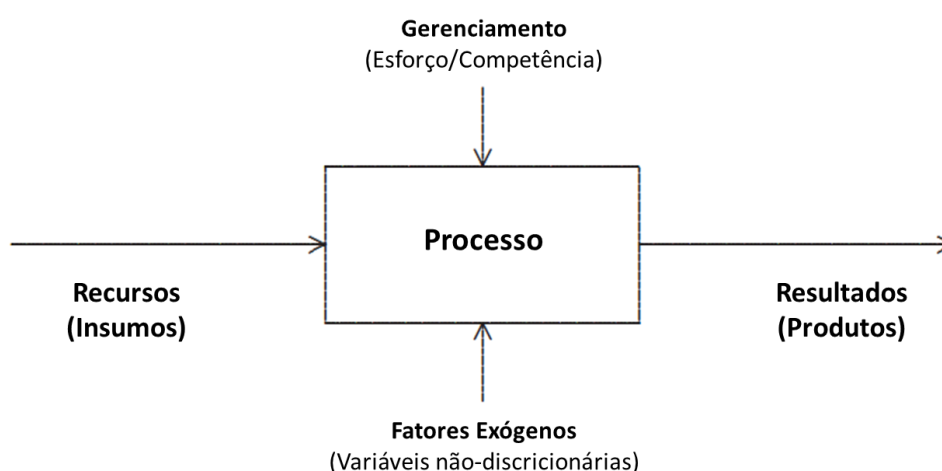


Figura 14. Visão sistêmica para avaliação da eficiência
Fonte: Bogetoft & Otto (2011).

Nesse sentido, um aspecto central que a avaliação de eficiência visa captar é o desempenho dos esforços empreendidos e do gerenciamento dos recursos, tecnologias e demais fatores envolvidos no processo produtivo. Para tanto, assim como proposto pelas teorias de Avaliação (Suchman, 1967; Scriven, 1991) e previsto por determinados modelos de avaliação em TD&E (e.g. Warr et al., 1970; Stufflebeam, 2003) realiza a comparação entre alternativas de alocação dos insumos, visando definir a estratégia ou tecnologia mais eficiente, seja em termos de melhores resultados ou menor consumo de recursos. A avaliação estabelece, assim, um *benchmark*, um referencial real ou hipotético, para a alocação ótima dos recursos na geração do máximo de resultados com os insumos disponíveis (Ferreira & Gomes, 2020).

Conforme Bogetoft e Otto (2011), a avaliação da eficiência visa elucidar aspectos gerenciais que potencialmente podem afetar a eficácia e, conseqüentemente, a efetividade dos processos em uma organização. Aplicado ao TD&E, a detecção de ineficiências alocativas em um sistema de treinamento pode indicar a necessidade de ajustes em suas atividades, reformulação de objetivos, cotejamento de estratégias alternativas ou mudanças na estrutura de recursos, de forma que se gerem economias ou que se ampliem os produtos e resultados de treinamento permitindo que se treine um maior número de indivíduos, ou que se forneçam melhores cursos, materiais e professores.

Muitos dos modelos de avaliação em TD&E se originaram no campo da Educação, tendo sido incorporados em razão de sua similaridade com relação aos objetivos de avaliação. Da mesma forma que desenvolveu esses modelos para avaliação da efetividade, o campo da Educação também desenvolveu uma vasta literatura orientada para a avaliação da eficiência de instituições educacionais (De Witte, & López-Torres, 2017; Johnes, Portela, & Thanassoulis, 2017). Dentre as metodologias de avaliação da eficiência educacional existentes, a mais difundida é o modelo de Análise Envoltória de Dados (*Data Envelopment Analysis* - DEA) (Emrouznejad & Yang, 2018).

O DEA foi desenvolvido na década de 1970 por Edward Rhodes com o propósito de se avaliar comparativamente a eficiência de escolas no âmbito de um programa público educacional. Para tanto, buscou estabelecer um método que fosse capaz de considerar múltiplos insumos e produtos os quais, para o projeto pioneiro em que foi desenvolvido, consistiam na aprendizagem de habilidades cognitivas, afetivas e psicomotoras pelos alunos, tendo por consideração dois tipos de insumos: horas de estudo em sala de aula; e tempo gasto pela mãe com leituras para o filho. A solução metodológica encontrada baseou-se na utilização de uma abordagem de programação matemática linear para o cálculo das relações ótimas entre insumos e produtos, orientando-se para a maximização dos resultados, o que permitia encontrar a

máxima eficiência possível, ou seja, quais as condições em que os alunos obtiveram o melhor desempenho tendo em vista o tempo de estudo (Barbosa & Fuchigami, 2018). A técnica foi, então, denominada Análise Envoltória de Dados e publicada no artigo seminal de Charnes, Cooper e Rhodes (1978).

Desde a publicação de seu artigo seminal, O DEA tem sido considerado uma das técnicas mais adequadas para a avaliação de eficiência em contextos que dependem da consideração de vários insumos e produtos simultaneamente, assim como situações em que tais variáveis não são facilmente mensuráveis (Afonso, 2007; Bretschneider et. al., 2005; Buleca & Mura, 2014; Emrouznejad et. al., 2014; Rosano-Peña, 2008; Smith & Street, 2005). Revisões de literatura específicas sobre modelagens baseadas no DEA podem ser encontradas em grande quantidade de campos, denotando a existência de um *corpus* teórico-metodológico relevante para avaliação da eficiência.

Além do campo da Educação (e.g. De Witte, & López-Torres, 2017) no qual se deu a aplicação pioneira seguida de grande difusão do DEA, também podem ser encontradas revisões acerca da aplicação desse modelo em temas como o de desenvolvimento humano (Mariano, Sobreiro & Rebelatto, 2015), eficiência energética (Mardani, et al., 2017), conservação ambiental (Song et al., 2012); atenção primária em saúde (Pelone et al., 2015); economia energética e ambiental (Mardani, et al., 2017), sustentabilidade ecológica organizacional (Zhou, Yi, Chang & Zhu, 2018), gestão de cadeias de suprimentos (Soheilrad et al., 2018), sistemas de transporte (Mahmoudi, Emrouznejad, Shetab-Boushehri & Hejazi, 2020), desenvolvimento sustentável (Tsaples & Papathanasiou, 2021), sistemas de saúde hospitalar (Kohl, Schoenfelder, Fügner & Brunner, 2019), desempenho de aeroportos (Iyer & Jain, 2019), setor bancário (Kumar & Singh, 2014), seleção de fornecedores em compras e aquisições (Dutta, Jaikumar, Arora, 2021) e indústrias florestais (Sowlati, 2005). A ampla utilização do DEA também é atestada por revisões do próprio campo (e.g. Emrouznejad & Yang, 2018), assim como em revisões que demonstram que tem sido o principal modelo para avaliação de eficiência e de qualidade do gasto no setor público (e.g. Sant'Anna et al., 2020).

Possivelmente, a difusão do DEA, enquanto um modelo sistêmico, se deve àquilo que almejavam os matemáticos quando desenvolveram a Teoria Geral de Sistemas na década de 1950: a integração da ciência sob um mesmo paradigma. Assim, o campo de TD&E, na medida em que possui grande proximidade junto ao campo da Educação e fundamenta-se firmemente no paradigma sistêmico, afigura-se como um campo em que a estruturação dos objetivos de avaliação pode possuir grande convergência com os pressupostos necessários à especificação de um modelo baseado na Análise Envoltória de Dados.

O próximo capítulo, portanto, aprofunda a discussão acerca da avaliação da eficiência e os pressupostos do DEA, tendo como objetivo a operacionalização do modelo de avaliação em TD&E proposto na presente seção. Visando estabelecer um paralelo com o campo da Educação, o capítulo busca ilustrar as aplicações do DEA em avaliação educacional, seus níveis, objetivos e variáveis de análise. O objetivo dessa integração reside na busca de subsídios para a posterior definição de variáveis que podem ser operacionalizadas no âmbito do TD&E.

2.3 Avaliação da eficiência

A avaliação da eficiência tem seus fundamentos em teorias do campo da Economia, cujas metodologias foram incorporadas pelas ciências organizacionais, sobretudo por intermédio do campo da Pesquisa Operacional, que tem como um de seus objetos de estudo a análise das possibilidades de produção de sistemas organizacionais (Forsund & Sarafoglou, 2002). Segundo Fávero & Belfiore (2012), a Pesquisa Operacional surge de maneira formal no contexto da Segunda Guerra Mundial (1939-1945) como um campo voltado para a solução de problemas de natureza logística, tática e de estratégia militar, envolvendo o desenvolvimento de modelos para decisão acerca da utilização mais eficaz dos recursos militares limitados. Para tanto, a Pesquisa Operacional lança mão da utilização de métodos científicos, baseados em modelos matemáticos, estatísticos e algoritmos computacionais para a tomada de decisões.

Sumariamente, os problemas de tomada de decisão são delineados em termos de funções matemáticas – denominadas função objetivo – e condicionantes – chamados de restrições – que visam determinar como alcançar um valor-alvo ou nível de qualidade para determinados objetivos em função das variáveis processuais e contextuais que integram o problema. A função objetivo pode visar maximizar um resultado ou minimizar um efeito indesejado (e.g. custos, erros, riscos). Desde o seu desenvolvimento, os conhecimentos da Pesquisa Operacional vêm sendo aplicados para a otimização de recursos em diversas áreas, tais como as de estratégia, *marketing*, finanças, logística e recursos humanos (Favero & Belfiore, 2012).

Com efeito, Suchman (1967), ao definir a avaliação da eficiência, apresenta o campo de Pesquisa Operacional como uma abordagem valiosa para essa perspectiva de avaliação em programas públicos. Segundo o autor, por meio da modelagem de um determinado programa, o propósito da Pesquisa Operacional visa determinar “quais cursos alternativos de ação são mais efetivos (ótimos) em relação a um conjunto pertinente de objetivos do tomador de decisão” (Suchman, 1967, p. 66). Dessa forma, o principal aspecto distintivo da avaliação da eficiência em perspectiva somativa é a avaliação do que foi realizado em relação às alternativas ótimas

do que poderia ter sido feito com os mesmos recursos disponibilizados (Barbosa & Fuchigami, 2018; Ferreira & Gomes, 2020; Lovell, 1993; Mariano, 2007).

Tal aspecto distingue-se e avança no que se refere a medidas tradicionalmente utilizadas como forma de avaliação e monitoramento da eficiência nas organizações, sobretudo no contexto da avaliação de programas e atividades do setor público (Brauchle & Schmidt, 2004; Buleca & Mura, 2014; Rosano-Peña, 2008). Por muito tempo, a perspectiva da eficiência nesse contexto baseou-se na avaliação de indicadores de desempenho, assumindo-se como eficiência o aumento dos resultados em indicadores de produtos e resultados (Ferguson, 2019, Manzoor, 2014, Smith & Street, 2005).

Tais medidas não permitem, porém, definir com precisão os parâmetros e alternativas de eficiência, eficácia e efetividade perante os quais o desempenho possa ser analisado. Sob a perspectiva da Pesquisa Operacional, tais indicadores relacionam-se à produtividade, a qual normalmente se operacionaliza por meio de uma razão entre um produto e seu(s) insumo(s) (e.g. a quantidade de alunos formados por professor, toneladas de cereal por hectare ou peças produzidas por funcionários em uma hora). Nesse campo, índices de produtividade compostos por medidas de produto e insumo são compreendidos como um fator de produtividade média parcial (Rosano-Peña, 2008). De uma maneira mais restrita, tais indicadores podem servir como indicadores de eficiência, posto que, quanto maior seus valores, maior será a eficiência do programa de treinamento, mas elas não avaliam um sistema produtivo como um todo, tampouco estabelece alternativas de alocação de recursos (Barbosa & Fuchigami, 2018; Lovell, 1993; Mariano, 2007; Rosano-Peña, 2008).

Conforme proposto pela literatura de Avaliação (e.g. Scriven, 1991; Suchman, 1967), a eficiência deve ser avaliada por meio da consideração das alternativas de alocação desses recursos, visando a otimização dos valores atribuídos a resultados pretendidos pelas atividades organizacionais (Manzoor, 2014; Sant'Ana et al., 2020; Rosano-Peña, 2008). Nesse sentido é que, apesar de manterem uma relação próxima no âmbito da literatura, produtividade e eficiência guardam uma importante distinção que, para que o campo de TD&E, pode proporcionar avanços para além da consideração de indicadores, como os observados, por exemplo, em Ford (1993) e Baraldí & Cifalinò (2015), de forma que ele incorpore a complexidade sistêmica inerente ao sistema de TD&E como implicado pelos modelos integrativos de avaliação (Abbad et al., 2020). Visando demonstrar as distinções entre esses dois conceitos e apresentar como a perspectiva da eficiência é tratada pelos campos da Economia e da Pesquisa Operacional, a próxima seção aprofunda a discussão acerca dessas definições.

2.3.1 O conceito de eficiência

A concepção mais comum de eficiência a define como uma medida baseada na razão entre os recursos utilizados e os resultados obtidos (Manzoor, 2014; Sant´Ana et al., 2020). Tal definição se aproxima do conceito mais técnico de eficiência, o qual a define como a melhor combinação possível de insumos e métodos (*inputs*) em um determinado processo produtivo para a geração do máximo de produtos (*outputs*) (Rosano-Peña, 2008). Dito de outra forma, a eficiência resulta da comparação entre o que foi produzido com determinada quantidade de recursos e o que poderia ter sido produzido com estes mesmos recursos (Barbosa & Fuchigami, 2018; Ferreira & Gomes, 2020). Assim, como demonstrado na Expressão (1), uma medida de eficiência determinará um valor adimensional de 0 a 1 (Mariano, 2007).

$$Eficiência = I / I_{max} \quad (1)$$

Em que:

I: Indicador de desempenho atual de um determinado sistema;

I_{max}: Máximo valor que o sistema pode alcançar nesse indicador

O conceito de produtividade, por sua vez, guarda uma relação direta com o conceito de eficiência, pois também consiste em uma relação entre insumos e produtos e ambos destinam-se à mensuração do desempenho de unidades produtivas (Lovell, 1993). A produtividade pode ser definida pela Expressão (2) (Mariano, 2007):

$$Produtividade = Y/X \quad (2)$$

Em que:

Y: *quantidade do produto*;

X: *quantidade do insumo*.

Como a produtividade é o resultado da divisão de duas quantidades diferentes, com suas respectivas unidades de medida, seu cálculo gera um indicador baseado em unidades de medida dos produtos e insumos (e.g. quilos por hora – kg/h), de forma que sua quantificação será diferente para cada caso. Por outro lado, uma medida de eficiência será uma medida geral e adimensional do sistema como um todo, de forma que permite a sua aplicação e interpretação da mesma maneira em todas as situações nas quais existam uma relação entre insumos e produtos (Mariano, 2007).

As diferenças entre os conceitos de eficiência e produtividade podem ser melhor demonstradas por meio de uma função de produção, oriunda da teoria microeconômica da produção (teoria da firma), que visa demonstrar a relação entre os produtos e insumos. A função de produção representada abaixo (Figura 15) e descrita pela equação (3) expressa a relação

entre a quantidade “X” de “n” fatores de produção (insumos) e a máxima quantidade “Q” de produtos ou serviços gerados a partir desses, pressupondo-se a eficiência perfeita (Ferreira & Gomes, 2020; Rosano-Peña, 2008).

$$Q_y = f_y(X_1, X_2 \dots X_n) \quad (3)$$

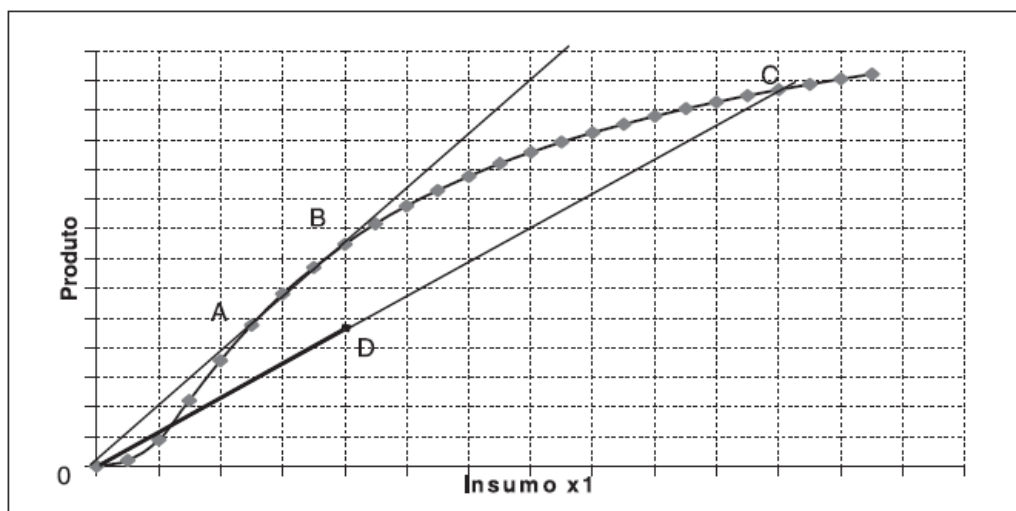


Figura 15. Função Produção

Fonte: Rosano-Peña (2008).

No gráfico, a curva pontuada representa a função de produção, chamada de fronteira de possibilidades de produção. A curva visa representar as relações mais eficientes possíveis entre insumos (X_n) e produtos (Q_y), sobre a qual encontram-se as firmas A, B e C, consideradas eficientes. A área abaixo dessa linha representa as possibilidades menos eficientes de produção, onde encontra-se, por exemplo, a firma D. Como pode-se verificar, o segmento \overline{OA} da função de produção representa uma elevação da curva pontilhada, designando retornos crescentes – em que o incremento à produção aumenta mais com a mesma proporção de incremento nos insumos – seguido de retornos constantes – dado pela direção retilínea da função – no segmento \overline{AB} e retornos decrescentes a partir de B, em que a curva sofre uma inflexão para baixo. Resgatando-se o conceito de produtividade como razão entre produtos e insumos, apesar de A, B e C serem eficientes, a produtividade das firmas A e B – representada pela inclinação da reta que as tangencia – é maior que a produtividade das firmas C e D, representada pela linha que cruza ambas essas firmas. Mesmo que C e D possuam a mesma produtividade, esta última, no entanto é ineficiente para a escala em que está operando. Para que D fosse considerada eficiente deveria aumentar sua produção até o ponto B ou reduzir a quantidade de insumos utilizados até próximo de A, de forma a se posicionar na fronteira de possibilidades eficientes de produção – as

alternativas perante a qual D é avaliada. Assim, matematicamente, caso as unidades elencadas estivessem sendo avaliadas apenas com base em seus indicadores de produtividade – ou seja, a simples razão entre insumos e produtos – se atribuiria o mesmo nível de desempenho para as firmas C e D, levando a uma avaliação incorreta, posto que C – perante às alternativas de produção - é mais eficiente do que D (Ferreira & Gomes, 2020; Rosano-Peña, 2008).

Por esse motivo, avaliar um sistema apenas com base em medidas e indicadores de produtividade pode levar a equívocos com relação à sua eficiência. Não obstante, indicadores de produtividade têm uma série de utilizações para avaliação e comparação de desempenho entre diferentes ciclos produtivos ou diferentes sistemas organizacionais (Lovell, 1993; Mariano, 2007). Porém, quando se trata da avaliação acerca das melhores decisões para alocação de recursos, os valores de determinado indicador de produtividade podem não explicar se uma firma é eficiente ou não. Os índices de produtividade parcial são mais apropriados para avaliação de metas e objetivos, ou seja, indicadores eficácia, por exemplo (Rosano-Peña, 2008).

Um sistema produtivo real normalmente é muito mais complexo e envolve múltiplos insumos e produtos, além de aspectos contextuais que podem influenciar em seus processos e resultados. Em vista disso, para se considerar uma maior quantidade de fatores de produção relevantes, foi desenvolvido o conceito de produtividade total (Lovell, 1993; Mariano, 2007). O cálculo da produtividade total ilustra a relação entre o total de produtos e o total de insumos. Como não assume a homogeneidade de insumos e produtos, atribui-lhes pesos com relação à sua utilidade (importância) para o resultado do processo produtivo. Por exemplo, os produtos podem ter seu peso definido pelo seu preço de venda, de forma que certos produtos são mais valiosos que outros, sendo o mesmo aplicado aos insumos. A utilidade do produto também pode ser entendida como a taxa de troca desse produto, enquanto a utilidade do insumo pode ser denominada como taxa de substituição desse insumo (Mariano, 2007). A razão final ocorre entre um produto e um insumo virtuais, resultantes da agregação ponderada desses elementos individuais de produção. Assim, a Produtividade Total pode ser representada pela Expressão (4) (Mariano, 2007):

$$Produtividade = \frac{u_1 \cdot y_1 + u_2 \cdot y_2 + u_3 \cdot y_3 \dots + u_i \cdot y_i}{v_1 \cdot x_1 + v_2 \cdot x_2 + v_3 \cdot x_3 \dots + v_j \cdot x_j} = \frac{Y_V}{X_V} \quad (4)$$

Em que:

Y_V : Medida virtual do produto;

X_V : Medida virtual do insumo.

u_i : Utilidade (peso) do produto i

y_i : Quantidade do produto i ;

v_j : Utilidade (peso) do insumo j ;

x_i : Quantidade do insumo j ;

Sob essa noção de produtividade total, a eficiência produtiva de uma unidade, seguindo o mesmo conceito de eficiência da Expressão (1), seria dada pela Expressão (5). A expressão demonstra matematicamente o conceito da eficiência, que se define pela produtividade total observada, considerados todos os insumos e produtos ponderados, em relação à produtividade total máxima que os sistema poderia ter alcançado (Mariano, 2007):

$$\text{Eficiência Produtiva} = P / P_{\max} \quad (5)$$

Em que:

P : Produtividade de um determinado sistema;

P_{\max} : Produtividade máxima que o sistema pode alcançar.

Quando os pesos de insumos e produtos são conhecidos, por exemplo os preços, salários ou outra medida de valor, o cálculo da produtividade pode ser facilmente realizado (Barbosa & Fuchigami, 2018). Mas, muitas vezes esses pesos e preços não estão disponíveis, não são determináveis ou quando são, não representam o real valor e importância do produto ou do insumo (Mariano, 2007). A indeterminação dos pesos ocorre em situações em que não há como determinar o valor de mercado dos fatores de produção, ou seja, eles não podem ser adequadamente analisados sob uma perspectiva monetária.

Este seria precisamente o caso de serviços intangíveis como serviços públicos de saúde, educação, segurança, entre outros. Tal situação também ocorre em casos de mercados imperfeitos em condições de monopólio como as compras e vendas públicas (Rosano-Peña, 2008). Um exemplo seria um hospital que tenha por insumos o quantitativo de médicos, enfermeiros e equipamentos e como produtos o tempo de permanência de pacientes e a taxa de mortalidade (Barbosa & Fuchigami, 2018). A dificuldade ou a inviabilidade de se determinar o valor desses produtos, ou de quaisquer outros fatores similarmente intangíveis, gera dificuldades para o cálculo da produtividade real e máxima e, conseqüentemente da eficiência desses sistemas produtivos (Mariano, 2007; Rosano-Peña, 2008).

Naturalmente, não apenas em contextos do serviço público essa dificuldade de determinação da importância relativa dos fatores é observada. As revisões sobre avaliação da eficiência em campos como os da Educação, Conservação Ambiental e Saúde sugerem que há um esforço da literatura nesses campos em se avaliar a eficiência relacionada a sistemas com fatores intangíveis, tais como a aprendizagem de alunos, a preservação ambiental e o bem-estar de pacientes. Conforme observado na revisão de literatura do campo de TD&E, suas atividades

também são compreendidas como orientadas para resultados majoritariamente intangíveis, perpassando pela aprendizagem individual, de equipes, organizações e sociedade (Aguinis & Kraiger, 2009). Nesse ínterim, as semelhanças entre as questões e dificuldades que surgem para a avaliação da eficiência em serviços públicos também podem ser um fator observado no campo de TD&E.

Diante de questões dessa natureza, formularam-se no campo da Economia algumas definições específicas de eficiência, assim como técnicas e métodos para sua avaliação que visavam apoiar a tomada de decisão de alocação de recursos nesses contextos de indeterminação de preços, sobretudo no setor público (Manzoor, 2014; Rosano-Peña, 2008; Smith & Street, 2005). A solução encontrada e desenvolvida pelos trabalhos seminais de Farrel (1957) consiste na decomposição da eficiência em dimensões específicas, de forma a se permitir a sua avaliação prescindindo-se da definição apriorística dos preços (coeficientes de peso) para os fatores de produção.

2.3.2 Tipos de eficiência

A eficiência é um conceito geral que pode ser aplicado para uma série de situações e devido à sua complexidade, foi desdobrado em tipos mais específicos com a finalidade de se estender a diferentes objetos de pesquisa e avaliação. As principais distinções clássicas de tipos de eficiência foram concebidas pelo trabalho seminal de Farrell (1957), que distinguiu problemas de eficiência produtiva, que envolvem os aspectos estritamente técnicos do processo produtivo, e problemas de eficiência econômica em que se consideram os custos e retornos da produção (Mariano, 2007; Rosano-Peña, 2008).

A eficiência produtiva de uma unidade é definida pela relação entre os insumos e produtos e possui dois componentes: a eficiência técnica e a eficiência de escala (Lovell, 1993). Diz-se de um sistema que este possui eficiência técnica – ou eficiência técnica pura – caso realize a melhor conversão de insumos em produtos, sem qualquer desperdício, de forma que haja pleno emprego dos fatores de produção. Nesse sentido, relaciona-se à forma como os recursos são utilizados, à qualidade do processo produtivo, à qualificação da força de trabalho e às técnicas e tecnologias adotadas (Ferreira & Gomes, 2020; Mariano, 2007; Rosano-Peña, 2008).

Por outro lado, a escala da produção também pode afetar a sua eficiência, de forma que para determinado nível de produtividade, há uma escala ótima de conversão de certa quantidade de insumos em produtos. Sendo assim, diz-se que uma unidade produtiva possui eficiência de escala caso atinja o seu tamanho ótimo, de forma que quaisquer alterações, para mais ou para menos, tornam a unidade ineficiente. Logo, a eficiência produtiva total, também chamada de

eficiência técnica global é alcançada quando uma unidade possui tanto eficiência técnica pura quanto eficiência de escala, os quais referem-se aos aspectos técnicos e físicos do sistema produtivo (Mariano, 2007; Rosano-Peña, 2008).

Para além dos aspectos puramente técnicos de produção, há uma segunda categoria de fatores que determina um novo tipo de eficiência, a alocativa. Uma unidade com perfeita eficiência produtiva, por exemplo, pode não possuir um bom desempenho caso as combinações – ou mix – de insumos e produtos, considerando-se o seu preço ou utilidade, não sejam a melhor combinação possível para minimizar os custos e maximizar os resultados ou retorno do sistema produtivo. Em outras palavras, diz-se que uma unidade possui eficiência alocativa quando consegue definir uma combinação de insumos e produtos que, respectivamente, minimizam os custos e maximizam os retornos ou a utilidade de seus resultados (Mariano, 2007; Rosano-Peña, 2008).

A eficiência alocativa determina, em conjunto com a eficiência técnica global, o conceito de eficiência econômica, que seria o estado ideal de operação do sistema produtivo, a sua versão mais viável (Mariano, 2007). Para Farrell (1957) a eficiência econômica pode também ser chamada, portanto, de eficiência global, produto das eficiências técnica e alocativa. Como exemplo ilustrativo, uma unidade de TD&E tecnicamente eficiente, mas cujos cursos não são demandados – não supre nenhuma necessidade – possui eficiência produtiva, mas não possui eficiência econômica por ausência de eficiência alocativa, haja vista que poderia estar empregando os mesmos recursos em ações educacionais que realmente atendessem necessidades e tivesse utilidade para os indivíduos, de forma a proporcionar a sua participação nessas atividades. A relação entre esses conceitos de eficiência pode ser representada conforme a Figura 16.

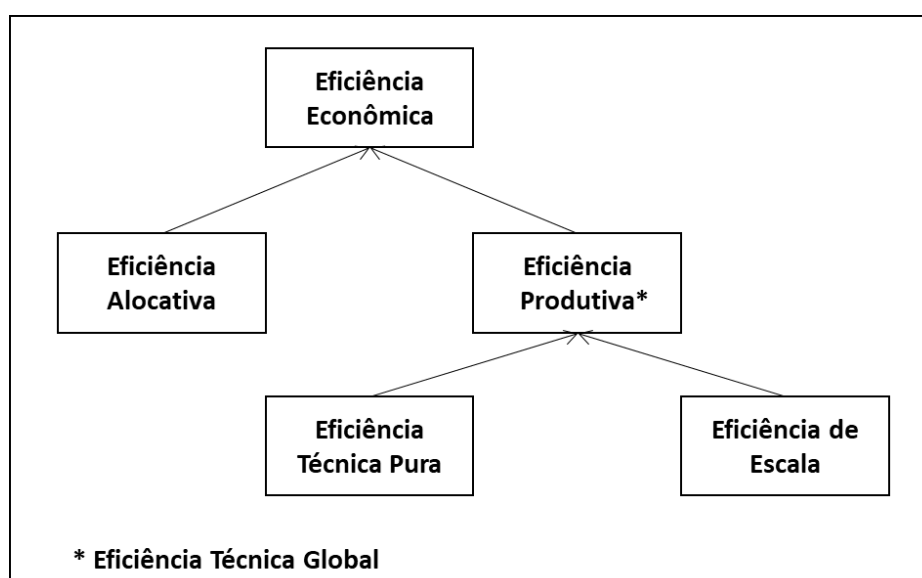


Figura 16. Tipos de eficiência
Fonte: Mariano (2007).

A Tabela 3, por sua vez, sumariza os tipos de eficiência, considerando também variações nas suas denominações.

Tabela 3
Tipos de Eficiência

Tipo de Eficiência	Característica
Econômica ou Geral	Capacidade de evitar a escassez e maximizar a satisfação humana. Resulta do produto entre as eficiências alocativa e produtiva.
Alocativa	Capacidade de alocar os recursos econômicos da melhor maneira possível incorrendo em menores custos e maiores resultados.
Produtiva, ou Técnica Global	Capacidade de gerar o máximo de produtos com dado montante de insumos ou de consumir menos insumos para dada quantidade de produtos.
Técnica Pura	Parcela da eficiência produtiva relacionada a fatores de ordem puramente técnica que afetam a taxa de conversão de insumos em produtos.
De Escala	Parcela da eficiência relacionada à melhor escala para determinado nível de produção, advinda da variação dos retornos na relação entre insumos e produtos.

Nota. Fonte: Elaboração própria com base em Mariano (2007).

Sendo assim, a avaliação de eficiência pode ser simplificada quando se torna possível decompô-la nas dimensões específicas acima descritas. Tais conceitos são úteis para determinadas avaliações que tenham objetivos intencionalmente mais específicos. Por exemplo, pode ser desejável avaliar um equipamento ou um processo produtivo tendo-se em consideração apenas a sua eficiência técnica pura – ser mais rápido, mais econômico – ignorando-se eventuais questões de escala ou de preço dos insumos (alocativa). Tal avaliação permite comparar tecnicamente equipamentos e processos entre firmas de diferentes portes (escala) ou regiões (alocativa), evitando-se a consideração dessas variáveis contextuais seja para que se isole esses efeitos, ou por serem questões que eventualmente fujam ao controle do tomador de decisão (Lovell, 1993).

Aplicado ao TD&E, essa decomposição traz a possibilidade de tornar mais específico o conceito de eficiência para o campo. Conforme demonstrado anteriormente, o critério da eficiência em TD&E visa assegurar “o máximo de resultados com o mínimo de fontes humanas, recursos financeiros e materiais” (Abbad et al., 2020, p. 227). Verifica-se que tal definição pode ser encontrada da mesma forma na Tabela 3 como o que é definido em Economia e Pesquisa Operacional como a “Eficiência Técnica Global”.

A vantagem dessa especificação é que os modelos de avaliação de eficiência técnica global serão capazes de estabelecer indicadores atribuíveis especificamente à eficiência técnica pura dos processos de planejamento, execução e aplicação de tecnologia instrucional de uma organização. Sob o ponto de vista da escala, poderá determinar por exemplo, qual o melhor nível de retorno para a quantidade de treinamentos realizados ou para os valores investidos, ou

seja, até que ponto novos investimentos ou novos treinamentos são capazes de gerar um incremento adequado nos resultados, permitindo o dimensionamento do Sistema de TD&E.

Com isso, conforme compreensão apresentada por Bogetoft & Otto (2011) a avaliação da eficiência técnica em TD&E refletirá as propriedades latentes e intrínsecas do Sistema de Treinamento relativas à sua capacidade de transformação de insumos em produtos e resultados. Tais indicadores de eficiência se referirão tão somente ao processo técnico de TD&E, isolando-o de possíveis fatores contextuais. Posteriormente, tais fatores contextuais podem ser objeto de avaliação por meio de técnicas estatísticas que visem estabelecer seu impacto na eficiência técnica do Sistema. Conclusivamente, torna-se possível saber a fonte das ineficiências – técnicas, de escala ou ambientais – de forma a oferecer alternativas de aperfeiçoamento tanto em nível de programa, quanto em nível organizacional, sendo esse, precisamente, um dos objetivos das avaliações (Mark et al., 2000).

A partir do trabalho de Farrell (1957), portanto, novas técnicas de análise tornam possível o cálculo da eficiência de maneira decomposta, permitindo análises em diferentes níveis para determinação da eficiência de unidades produtivas complexas, com múltiplos insumos, produtos e permitindo a consideração progressiva de diferentes fatores, tais como preços, escala e contexto. A subseção seguinte apresenta as abordagens para análise da eficiência, oriundas dos modelos de eficiência de Farrell (1957), tendo por ênfase a avaliação da eficiência técnica global, mais próxima do conceito de eficiência utilizado no campo de TD&E.

2.3.3 Métodos de Avaliação da Eficiência

Na medida que uma avaliação se destina a determinar o quanto um determinado programa ou atividade organizacional alcançou um resultado ou efeito esperado, uma das grandes questões para a realização de uma avaliação da eficiência reside em como estabelecer um parâmetro perante o qual as unidades produtivas sejam comparadas. Conforme discutido anteriormente, diferentemente das avaliações de indicadores de produtividade, as avaliações de eficiência se propõem a estabelecer alternativas ótimas como parâmetro para determinação da eficiência do sistema produtivo. Para tanto, devem buscar estabelecer quais são esses parâmetros ótimos de eficiência, os quais, nessa literatura, são chamados de *benchmarks* que representam as melhores práticas.

Para solucionar essa questão, os modelos baseados no conceito de eficiência de Farrel (1957) buscaram o desenvolvimento de medidas de eficiência com base em comparações empíricas entre unidades produtivas, chamadas nesse campo de *Decision Making Units* – DMUs (em tradução livre, unidades de tomada de decisão). Com isso, as avaliações se detêm

na caracterização de uma amostra de DMUs e na aplicação de métodos e técnicas para determinação de quais as alternativas mais eficientes de produção estão sendo observadas naquele conjunto específico de DMUs. As metodologias descrevem, assim, a fronteira eficiente de possibilidades de produção de forma que as unidades são, então, comparadas com relação à fronteira ótima, na qual se localizam os *benchmarks*, sendo esse o parâmetro por meio do qual as unidades são avaliadas e a forma por meio da qual são calculadas as magnitudes de suas ineficiências (Barbosa & Fuchigami, 2018; Ferreira & Gomes, 2020; Rosano-Peña, 2008; Smith & Street, 2005).

A avaliação parte de um modelo genérico para representação das DMUs com a definição dos insumos e produtos relevantes para análise de sua eficiência (Figura 17). Nesse sentido, surge a prerrogativa de que as unidades produtivas analisadas são homogêneas, no sentido de atuarem sob as mesmas condições de mercado, realizar as mesmas tarefas e utilizar os mesmos insumos para produção dos mesmos produtos, para que sejam efetivamente comparáveis (Rosano-Peña, 2008; Mariano, 2007). As DMUs são qualquer tipo de organização ou unidade autônoma que realize um processo de transformação de “n” insumos (*inputs*) em “n” produtos (*outputs*), os quais servem de critério para avaliação da sua eficiência (Ferreira & Gomes, 2020).



Figura 17. Representação de uma DMU

Fonte: Elaboração própria.

De forma geral, para a análise da eficiência não são necessárias a descrição e a consideração de variáveis intermediárias dos processos que ocorrem dentro do sistema produtivo para a conversão de insumos em produtos, sendo esse modelo tradicional também chamado por esse motivo de “caixa-preta” (*black-box*) (Cook et.al., 2010; Ferreira & Gomes, 2020; Yang et. al., 2018). Posteriormente, como será demonstrado adiante, extensões metodológicas podem ser utilizadas para tornar o modelo mais amplo, com a consideração de fluxos internos de processos, assim como fatores interferentes.

Tendo por base as pesquisas em produtividade, tradicionalmente existem duas classes de metodologias para o cálculo da eficiência: os métodos paramétricos e os métodos não-paramétricos (Afonso, 2007; Mariano, 2007; Rosano-Peña, 2008; Smith & Street, 2005). Os métodos chamados paramétricos, tais como as Funções Determinísticas e a Análise de

Fronteiras Estocásticas – SFA (*Stochastic Frontier Analysis*), utilizam métodos estatísticos, como a regressão multivariada, para estimar uma função produção a partir de um conjunto de insumos e produtos e avaliar as variações (ineficiências) em relação à tendência central (Afonso, 2007; Mariano, 2007; Rosano-Peña,2008; Smith & Street, 2005). Os métodos não-paramétricos, tais como a Análise Envoltória de Dados – DEA (*Data Envelopment Analysis*) e a Análise de Livre Disponibilidade – FDH (*Free Disposable Hulls*), estimam uma fronteira de possibilidades de produção que determina o nível ótimo de produção relativo a quantidades e combinações de insumos (Afonso, 2007; Mariano, 2007; Rosano-Peña,2008; Smith & Street, 2005). Cada uma dessas técnicas possui características, modelos, pressupostos e perspectivas próprios que devem se ajustar às condições e objetivos da análise (Mariano, 2007).

O objetivo desses métodos é realizar uma comparação entre o nível de desempenho de uma DMU e o desempenho que esta poderia ter alcançado, tendo em vista os resultados obtidos pela demais DMUs da amostra. Assim, permite definir a possível redução de insumos, mantidos os níveis de produção, ou inferir-se também qual o aumento na produção com o mesmo nível de insumos, caso esses fossem utilizados de forma mais eficiente (Afonso, 2007). A mensuração dessa lacuna determina a ineficiência da unidade produtiva em relação à fronteira eficiente calculada com os dados da amostra. Tendo em vista a ausência de preços para definição dos pesos dos insumos e produtos, normalmente o que os métodos avaliam é a eficiência técnica global. A Figura 18 **Erro! Fonte de referência não encontrada.** apresenta um resumo dos citados métodos e como estão relacionados.

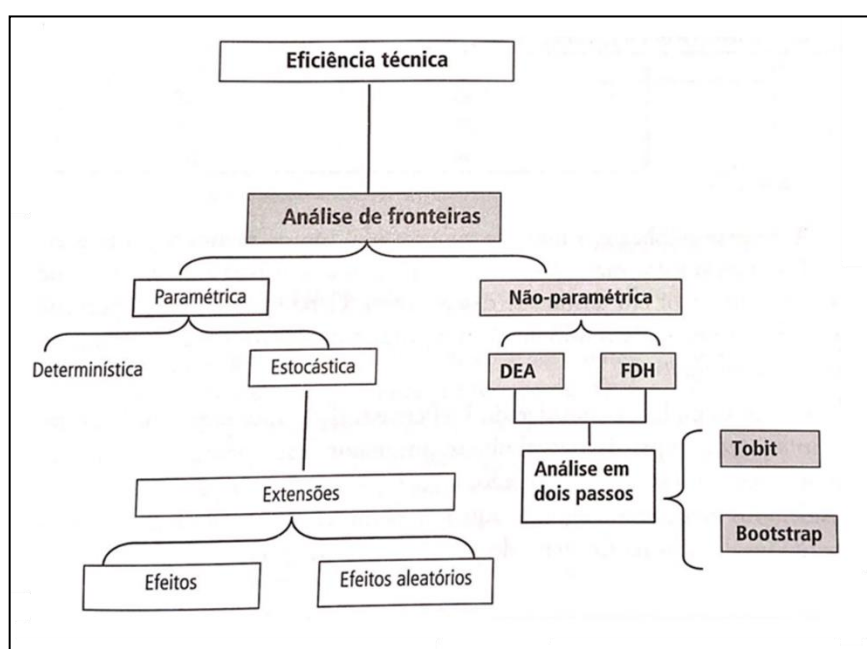


Figura 18. Métodos de Análise de Eficiência

Fonte: Afonso (2007).

Sumariamente, os métodos paramétricos são similares às análises estatísticas tradicionais. Baseiam-se na análise de regressão e por isso exigem a definição prévia da relação teórica entre insumos e produtos (Rosano-Peña, 2008). Os métodos paramétricos são considerados mais tradicionais (Rosano-Peña, 2008), porém têm sido criticados devido às suas propriedades inerentes às ferramentas estatísticas utilizadas (Smith & Street, 2005). O principal aspecto de crítica refere-se ao fato de que, por basear-se em uma regressão, a função produção é uma predição probabilística da média do desempenho da amostra (Rosano-Peña, 2008) e por isso, representa uma estimativa imprecisa da eficiência real para as unidades analisadas, posto que se baseia em uma projeção e não em uma fronteira empírica real.

Em contraste à abordagem paramétrica, os métodos não-paramétricos são desenvolvidos sem a utilização de inferências estatísticas, ou seja, sem se basear em medidas de tendência central, testes e estimações de coeficientes ou pressupostos e aspectos formais de análises de regressão (Afonso, 2007; Ferreira & Gomes, 2020). As abordagens não-paramétrica se baseiam naquilo que se apresentou como os métodos tradicionalmente utilizado em Pesquisa Operacional de programação matemática de otimização. Com isso, a análise prescinde da determinação das relações funcionais entre os insumos e produtos (Ferreira & Gomes, 2007; Rosano-Peña, 2008) e busca determinar uma curva de possibilidades de produção com base apenas nas observações realizadas (Smith & Street, 2005), ao contrário da abordagem paramétrica que estabelece uma projeção oriunda da tendência central a partir do comportamento de todas as organizações observadas (Barbosa & Fuchigami, 2018).

Em contrapartida, a natureza não estatística dos modelos não-paramétricos torna-os determinísticos, susceptíveis às observações extremas e aos erros de medida. Além disso, por prescindirem de pressupostos estatísticos, essas técnicas possuem limites para a realização de inferências acerca das causas da eficiência (Ferreira & Gomes, 2020). Tendo em vista essas limitações, conforme demonstrado na Figura 18, análises chamadas de análises em dois passos podem ser realizadas, delineando os chamados modelos de dois estágios.

A análise em dois estágios perpassa, primeiramente, pela análise da eficiência. No segundo estágio, por sua vez, uma série de extensões metodológicas complementares com utilização de métodos estatísticos pode ser utilizada para apoiar a interpretação dos dados. Segundo Afonso (2007) e Boueri et al. (2007), um dos principais métodos complementares refere-se à aplicação de técnicas de estimação de intervalos de confiança por meio de técnicas de *bootstrapping* para que se estime a significância das diferenças de eficiência observadas, o que possibilita atenuar o caráter determinístico dos modelos.

Uma segunda técnica abrangida pelo segundo estágio consiste na utilização de regressões para análise da influência de fatores interferentes – chamados de fatores exógenos ou variáveis não-discrecionárias – nas eficiências observadas. Tendo em vista que os índices de eficiência variam de 0 a 1 são utilizados os métodos de regressão “Tobit”, o qual se aplica a variáveis “censuradas”, ou seja, que possuem uma amplitude limitada de valores que pode assumir, como é o caso dos indicadores de eficiência (Afonso, 2007; Boueri et al., 2007; Ferreira & Gomes, 2020). A realização dessa análise é fundamental para que se integre a influência do contexto na análise da eficiência.

Dentre as metodologias não-paramétricas para avaliação da eficiência, o Modelo DEA destaca-se por possuir maior flexibilidade aliada à maior precisão e rigor. Os modelos DEA permitem a decomposição da eficiência produtiva em suas eficiências técnica e de escala, podendo comportar variáveis numéricas, categóricas, instrumentais ou de decisão, e variáveis exógenas como medida de consideração do contexto. Além disso, modelos mais modernos permitem a análise longitudinal de dados em painel, tais como os modelos baseados no cálculo do índice Malmquist, o qual mensura a evolução da fronteira produtiva e a eficiência de unidades organizacionais (Ferreira & Gomes, 2020). Por fim, a combinação de modelos de análise encadeados, permitiu a emergência de modelos de DEA em rede, que visam superar as limitações dos modelos de “caixa-preta” ao delinear a análise como uma rede de etapas produtivas (Ferreira & Gomes, 2020; Kao, Chiu & Ko, 2014).

O *Network* DEA (N-DEA) ou DEA em rede caracteriza-se pela representação de um processo produtivo por meio de uma série de subprocessos interligados (Kao, 2017). A interação entre os processos sequenciais parte, assim como nos modelos tradicionais, da definição dos insumos iniciais (X) e dos produtos finais (Y). No entanto, interliga diferentes partes de um processo e a permite a consideração de variáveis intermediárias (Z) que são, ao mesmo tempo, produtos dos primeiros estágios e entradas para os estágios seguintes. Cook et. al. (2010) argumentam que os modelos em rede podem representar sistemas abertos, que admitem a entrada ou saída de recursos em quaisquer etapas do processo produtivo (Figura 19), ou sistemas fechados, em que todos os recursos são produzidos e consumidos dentro do próprio processo sem interação com o ambiente, salvo pelas entradas iniciais e saídas finais.

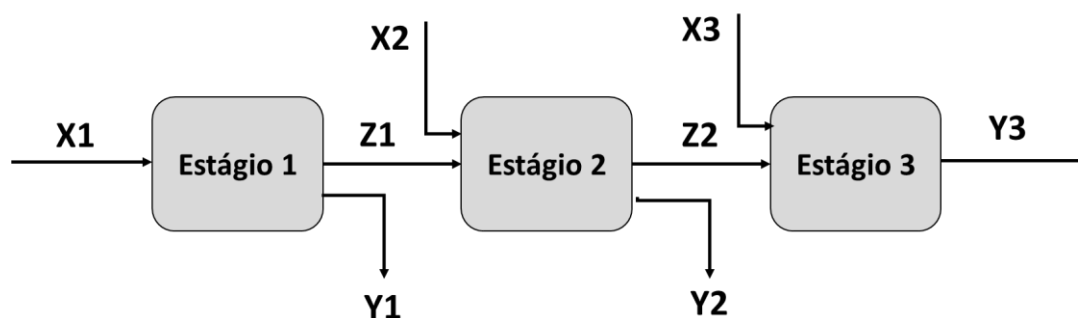


Figura 19. Representação de uma DMU em série de três estágios

Nota. Fonte: Cook et. al. (2010).

Nesse sentido, tendo em vista a clara perspectiva sistêmica do processo de treinamento e sua integração com o ambiente organizacional (Borges-Andrade & Abbad, 1991), assim como a existência de variáveis de diferentes naturezas implicadas em um processo de TD&E, o DEA enquanto método de avaliação da eficiência se configura compatível com os pressupostos teóricos e com as características do campo de TD&E. As extensões metodológicas para o DEA, tais como os Modelos em Dois Estágios e o N-DEA permitem considerar variáveis exógenas e operacionalizar as múltiplas etapas da estrutura básica do sistema de TD&E, que podem ser compreendidas, para fins de modelos de avaliação, sob uma perspectiva de contexto-insumo-processo-produto (Abbad et al., 2020).

Cabe salientar que os métodos de avaliação da eficiência, sobretudo o DEA, determinam os indicadores de eficiência tão somente com base nos dados da amostra de DMUs, tendo por consideração os insumos e produtos definidos pelo avaliador. Com isso não se permite a inferência e a generalização acerca das relações entre os fatores de produção para criação de uma teoria específica de eficiência naquele contexto. Pesquisas científicas posteriores, por outro lado, podem com desenhos experimentais e controles adequados, examinar essas relações de forma mais específica. Possivelmente, o acúmulo de estudos de eficiência de Farrel (1957), realizado por revisões de literatura, torna-se uma fonte para identificação de padrões de regularidade nas relações entre insumos e produtos que tenham sido operacionalizados de maneira similar por múltiplos estudos. Ainda assim, tais ferramentas de avaliação permitem, *a priori*, conclusões apenas dentro do contexto da amostra analisada e não estabelecem um padrão de eficiência para outros contextos. A mudança da amostra de DMUs, por exemplo, ensejará uma nova análise e, possivelmente, diferentes resultados de eficiência.

Sendo a avaliação o propósito do DEA, esse foi originalmente desenvolvido para a avaliação educacional com o intuito preciso de abordar problemas típicos de indeterminação de preços dos fatores de produção como os encontrados tipicamente no setor público. Como

expresso por seus autores, estes visavam “o desenvolvimento de medidas de eficiência na tomada de decisões, com especial referência à sua possível aplicação na avaliação de programas públicos” (Charnes, Cooper & Rhodes, 1978, p. 429). Segundo a revisão realizada por Emrouznejad & Yang (2018), marcando 40 anos desde a concepção do Modelo DEA por Charnes, Cooper & Rhodes (1978), este tem sido utilizado por um número crescente de pesquisadores, sendo amplamente aplicado tanto no setor público quanto privado. Nesse ínterim, novos modelos foram propostos, sendo que a próxima seção detalhará as funções matemáticas e os tipos de análises e resultados abrangidos pelos dois principais modelos de DEA, os quais são a base para avaliações e pesquisas de diferentes campos.

2.3.4 *Análise Envoltória de Dados*

O desenvolvimento da Análise Envoltória de Dados (DEA) se deu a partir da necessidade de ampliar as contribuições originalmente realizadas pelos modelos de mensuração da eficiência desenvolvidos por Farrel que presumiam a análise para um único produto (*output*), de tal forma que não havia suficiente desenvolvimento teórico para um modelo com múltiplos produtos (Cooper, Seiford & Zhu, 2011). Para resolver esta questão, Charnes, Cooper e Rhodes (1978) criaram o conceito de *input* e *output* virtuais e um método de programação linear para o cálculo de seus pesos. Consoante às abordagens da Pesquisa Operacional, uma função objetivo é construída orientada para a maximização da relação *input-output*, o que permitia encontrar a máxima eficiência possível e, portanto, a definição de um indicador relativo de eficiência para cada DMU (Barbosa & Fuchigami, 2018).

O primeiro modelo DEA, portanto, é referido como modelo CCR, em referência às iniciais dos seus autores Charnes, Cooper e Rhodes (1978). Neste, a eficiência relativa das unidades é calculada por meio da razão entre a soma ponderada dos produtos e a soma ponderada dos insumos, adotando como restrição do problema que os resultados devem sempre retornar valores entre 0 e 1. Assim, a sua formulação [6] representa o problema de programação linear no qual se apresentam N unidades, as quais produzem m quantidades de produtos y , utilizando-se de n quantidades de insumo x . Considerando que uma unidade o produziu y_{ro} produtos com x_{io} insumos, o que se busca é a determinação dos melhores valores para os pesos v_i e u_r para cada insumo e produto respectivamente, de tal forma que a eficiência produtiva da unidade (h_o) seja maximizada, dentro de uma restrição de que todas as medidas de eficiência sejam iguais ou menores que 1.

$$\begin{aligned}
 \text{Max } h_o &= \frac{\sum_{r=1}^m u_r y_{ro}}{\sum_{i=1}^n v_i x_{io}} \\
 \text{Sujeito a} & \\
 \frac{\sum_{r=1}^m u_r y_{rj}}{\sum_{i=1}^n v_i x_{ij}} &\leq 1 \quad j = 1, \dots, o, \dots, N \\
 u_r, v_i &\geq 0 \quad r = 1, \dots, m; \quad i = 1, \dots, n
 \end{aligned}
 \tag{6}$$

Para aplicação em situações reais, há uma distinção relevante que deve ser realizada *a priori*, acerca do objetivo da avaliação e que influenciará na forma como o modelo é operacionalizado e como seus cálculos são realizados. Quando o objetivo da avaliação é avaliar a melhor utilização dos recursos, a eficiência deve refletir a economia desses recursos e orientar a avaliação da eficiência no sentido da redução do consumo, com a manutenção dos níveis de resultados. Para essa perspectiva de análise, diz-se que o modelo tem orientação para insumos. Nesse caso, os valores que retornem o valor máximo “1” serão as unidades consideradas eficientes e em relação as quais, os índices de eficiência das demais serão calculados, retornando valores menores que “1” dimensionando a diferença a ser superada, ou seja, uma unidade ineficiente com *score* 0,80 deverá ajustar seus insumos, reduzindo-os em uma ordem de 20% (Rosano-Peña, 2008).

Alternativamente, o problema pode ser reformulado para que se tenha como objetivo o aumento dos resultados, para um mesmo nível de insumos, como forma de se aumentar a eficiência. Nesse caso, se denominará a análise como orientada para os produtos. No caso, as unidades eficientes continuarão a apresentar o índice 1, mas as ineficientes terão um índice correspondente ao inverso do indicador calculado pela orientação aos insumos, de forma que no exemplo anterior será de $1,25 = 1/0,8$. Assim a unidade deveria aumentar os bens e serviços produzidos em 25% para ser eficiente, sendo as análises voltadas para o sentido da melhoria de resultados.

A distinção entre essas duas formas de análise tem importância pelo fato de que deve levar em consideração se o objetivo da análise é buscar a eficiência por economia de insumos ou se o que se deseja é aumento do desempenho para o sistema analisado (Barbosa & Fuchigami, 2018). Além disso, para alguns casos, os insumos ou os produtos em determinado sistema produtivo podem ser fixos, ou seja, não podem ser facilmente ajustados. Para o caso de unidades de treinamentos, por exemplo, a supor que todas possuam uma meta fixa de capacitar

100 trabalhadores com um mesmo curso, será mais eficiente aquela que o fizer com menos recursos (orçamento, tempo, etc.), sendo adequado um modelo orientado para insumos. Para outro caso, por exemplo, em que as unidades recebam um mesmo orçamento fixo (insumos) de um milhão, será mais eficiente aquela que conseguir alcançar melhores resultados com essa mesma quantidade definida de insumos. Cabe salientar que uma combinação desses pressupostos pode ser utilizada, caracterizando um modelo não-orientado, em que tanto insumos quanto resultados podem ser otimizados. Porém, o modelo CCR apresenta limitações para algumas situações, como será demonstrado adiante.

A característica fundamental do Modelo CCR é que este assume os retornos constantes de escala, de forma que a fronteira de possibilidades de produção sobre a qual se localizam as unidades eficientes não se estenderá sob a forma de uma isoquanta curva, mas como uma reta diagonal a partir da origem do gráfico e envolvendo, as demais unidades (Ferreira & Gomes, 2020). Dessa forma, por ser um modelo linear, o tipo de orientação não deve afetar os resultados de eficiência para as unidades, sendo apenas formas diferentes de se representar seus índices (Rosano-Peña, 2008). A Figura 20 ilustra como são representadas as duas orientações e suas diferenças:

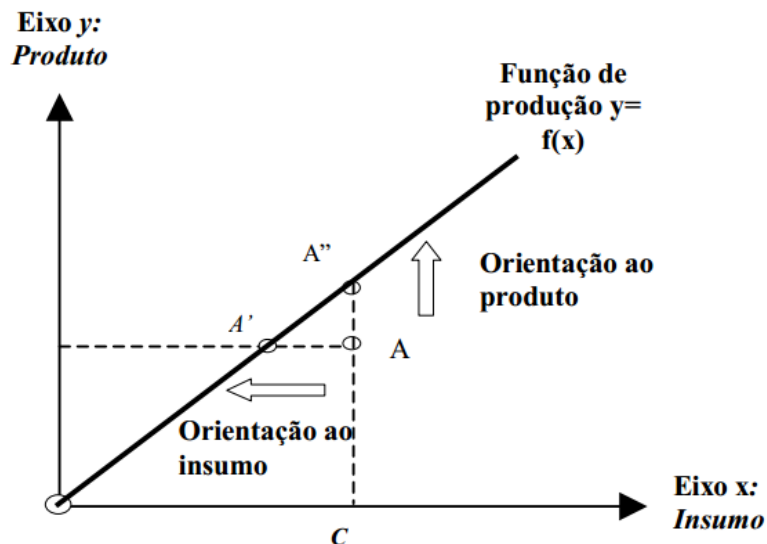


Figura 20. Orientação dos modelos matemáticos para os inputs e os outputs
Fonte: Mariano et al. (2006).

Conforme o gráfico, há duas alternativas para avaliação da unidade “A” que se apresenta como ineficiente. Caso se adote uma orientação ao produto, sua ineficiência é avaliada em relação ao ponto A” e seu objetivo é aumentar a produção, enquanto uma orientação ao insumo adotará como referência o ponto A’, denotando a redução nos insumos. Conforme mencionado, a orientação do modelo é definida aprioristicamente e define como a ineficiência será

mensurada e, conseqüentemente qual a extensão e em que variáveis o ajuste para otimização deve ser realizado.

A presunção de retornos constantes de escala (CRS – *Constant Returns to Scale*) assumida pelo CCR é plausível e empiricamente comprovada para uma grande parte dos setores produtivos tradicionais. Porém, para algumas atividades o processo produtivo pode apresentar oportunidades de economia de escala, ou ainda pode estar atuando com o nível de competência ideal, principalmente no setor público, de tal forma que podem ocorrer situações em que os retornos à escala sejam crescentes ou decrescentes (Rosano-Peña, 2008). Tendo em vista essa questão, seis anos mais tarde, outro modelo viria a ser considerado fundamental ao lado do modelo CCR, qual seja o BCC, apelidado, da mesma forma em referência aos seus criadores Banker, Charnes & Cooper (1984).

O modelo BCC é desenvolvido com a premissa de retornos variáveis à escala (VRS – *Variable Returns to Scale*) para avaliar a eficiência técnica. Desse modo, o modelo de proporcionalidade linear do CCR é substituído por uma representação convexa da fronteira de possibilidades produção, permitindo a identificação de retornos crescentes e decrescentes de escala. Essa diferença apresenta implicações importantes para a interpretação dos resultados e, não raro, as técnicas são aplicadas em conjunto para se comparar os resultados visando enriquecer a análise (Rosano-Peña, 2008). O BCC orientado para o produto podem ser dados pela seguinte formulação:

$$\text{Min } h_o = \sum_{r=1}^m v_r x_{ro} + v_o \quad [4]$$

Sujeito a

$$\sum_{i=1}^n u_i y_{io} = 1 \quad [4.1]$$

$$\sum_{r=1}^m u_i y_{rj} - \sum_{i=1}^n v_i x_{ij} - v_o \leq 0 \quad j = 1, \dots, o, \dots, N \quad [4.2]$$

$$u_r, v_i \geq 0 \quad r = 1, \dots, m; \quad i = 1, \dots, n \quad [4.3]$$

Quando se orienta para os insumos, a seguinte formulação é dada:

$$\text{Max } h_o = \sum_{r=1}^m u_r y_{ro} - u_o \quad [5]$$

Sujeito a

$$\sum_{i=1}^n v_i x_{io} = 1 \quad [5.1]$$

$$\sum_{r=1}^m u_r y_{rj} - \sum_{i=1}^n v_i x_{ij} - u_o \leq 0 \quad j = 1, \dots, o, \dots, N \quad [5.2]$$

$$u_r, v_i \geq 0 \quad r = 1, \dots, m; \quad i = 1, \dots, n \quad [5.3]$$

As diferenças dessas formulações com relação ao modelo CCR reside nos componentes adicionais V_o na primeira e u_o na segunda, os quais representam os retornos variáveis da escala que ajustam a formulação em relação ao modelo linear. Quando essas variáveis têm valor negativo, indicam retornos crescentes, quando positivo, retornos decrescentes; e por fim, retornos constantes de escalas quando têm valor nulo.

Em termos práticos, quando os retornos são crescentes o aumento nos insumos ocasiona um aumento desproporcionalmente maior nos produtos, denotando uma DMU operando abaixo da sua capacidade ótima. Para uma DMU eficiente, isso implica que ela deve elevar sua produção com ganhos de escala. Em situações de retorno constante, o aumento dos insumos ocasiona um aumento proporcional nos produtos, sendo o caso de uma DMU operando em capacidade ótima. Por fim, com retornos decrescentes, aumentos nos insumos ocasiona um aumento desproporcionalmente menor nos produtos, representando uma DMU que está operando acima da sua capacidade ótima. Sendo assim, o modelo BCC insere a consideração do porte das organizações, permitindo que os *benchmarks* para as organizações ineficientes sejam aqueles que mais se assemelham em termos de escala de produção (Mariano et al., 2006)

Quando consideradas as diferenças entre os modelos CCR e BCC, verifica-se que o formato convexo da fronteira eficiente gerado pelo BCC pode levar a diferenças com relação aos índices de eficiência dependendo da orientação adotada, contrariamente ao que ocorre com o modelo CCR. A fronteira criada por meio do método BCC, por sua natureza convexa capturará regiões com relações insumo-produto não lineares, representando as variações nos retornos de escala, e nas quais uma unidade produtiva pode apresentar uma distância (ineficiência) orientada para insumos diferente da distância orientada para produtos. A Figura 21 visa apresentar graficamente duas fronteiras de produção sobrepostas demonstrando as diferenças entre os modelos CCR e BCC.

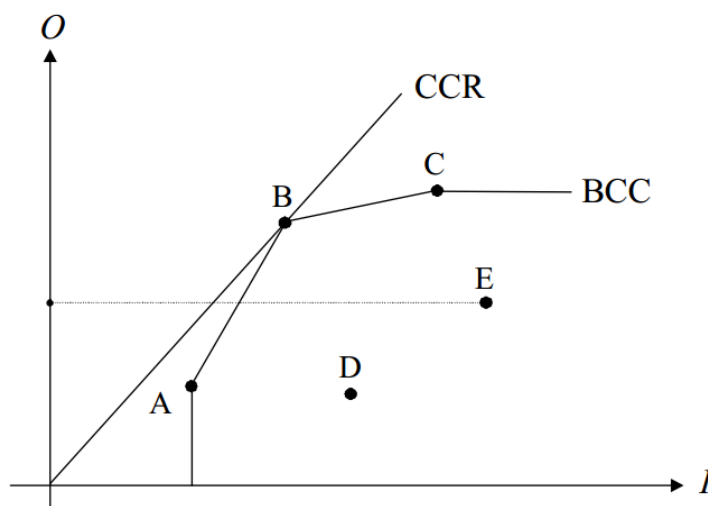


Figura 21. Comparação entre as fronteiras dos modelos BCC e CCR
 Fonte: Rosano-Peña (2008).

Conforme verifica-se na Figura 21, a ineficiência atribuída às unidades “D” e “E” no modelo BCC seria menor do que seria com base no modelo CCR, haja vista a maior proximidade da fronteira eficiente convexa. Além disso, no modelo BCC, a distância de “D” com relação à fronteira convexa é maior no sentido vertical (produtos) do que no sentido horizontal (insumos), sendo “A” a unidade de referência. A situação é contrária para “E”, em que “C” se torna a unidade de referência. Caso o modelo aplicado seja orientado para insumos, “D” seria relativamente mais ineficiente do que “E”, enquanto se forem considerados os produtos, a situação se inverte. Logo, os índices de eficiência seriam distintos entre os modelos DEA utilizados, e para o caso do modelo BCC, o tipo de orientação adotada influenciaria nos índices e no tipo de remediação para maximização da eficiência.

Os modelos apresentados, portanto, representam os dois modelos fundamentais de DEA. Tendo em vista o seu impacto nos resultados, a escolha do modelo e do tipo de orientação adotada terminam por configurar perspectivas de análise, com premissas e objetivos determinados. Não obstante, mais importante que essa definição da metodologia e da orientação é a importância de como são modelados os problemas de eficiência, isto é, quais as variáveis que devem ser consideradas, sua quantidade e qual a relevância dessas e das relações presumidas entre elas.

Conforme mencionado anteriormente, a avaliação da eficiência pública tem como principal obstáculo a definição e mensuração de suas variáveis de interesse e, sobretudo, a disponibilidade dos dados em todo o conjunto de DMUs a ser analisado. Assim, um dos pontos fundamentais da modelagem com Análise Envoltória de Dados é a definição correta das

variáveis de insumos e produtos que serão utilizadas para cálculo dos índices de eficiência das unidades analisadas (Ferreira & Gomes, 2020; Senra, Nanci, Mello & Meza, 2007).

A próxima seção tratará de alguns dos cuidados e limitações que devem ser levados em consideração para a definição das variáveis que constituirão um modelo DEA. Para apoiar a modelagem almejada no presente trabalho, de abordar a eficiência em TD&E, serão apresentadas, posteriormente, uma análise da literatura de avaliação da eficiência no campo da Educação e uma revisão de literatura para compreensão de como tais modelos podem ser operacionalizados e quais variáveis podem ser utilizadas especificamente em TD&E.

2.3.5 Modelagem e Seleção de Variáveis em DEA

Um dos pontos fundamentais da modelagem com Análise Envoltória de Dados (DEA – *Data Envelopment Analysis*) é a definição correta das variáveis de insumos e produtos que serão utilizadas para cálculo dos índices de eficiência das unidades analisadas (Ferreira & Gomes, 2020; Senra et. al., 2007). O DEA possui como uma de suas virtudes a capacidade de avaliar a eficiência considerando múltiplos insumos e produtos, de forma a melhor caracterizar a complexidade do sistema produtivo em questão. Por outro lado, essa vantagem encontra limites em si mesma, posto que, como explicam Ferreira e Gomes (2020) e Rosano-Peña (2008), a adição indiscriminada de variáveis tenderá a aumentar gradativamente os índices de eficiência devido ao aumento das possibilidades de utilização eficiente de recursos, o que tornará o modelo DEA incapaz de realizar a distinção e classificação do desempenho das DMUs (Ferreira & Gomes, 2020).

Além disso, segundo Rosano-Peña (2008), para que se tenha uma análise robusta, quanto maior o número de variáveis abrangidas por um modelo, maior a quantidade de DMUs que devem ser abordadas e conseqüentemente, maior o volume de dados a ser coletado e analisado. Uma recomendação comum na literatura é de que, para que haja confiabilidade dos resultados, o número de DMUs incorporadas na análise deve corresponder a cinco vezes a quantidade de variáveis, consideradas tanto as de insumo quanto de produto (Rosano-Peña, 2008). Dessa forma, diante de um conjunto de variáveis constituintes dos processos realizados pelas DMUs, deve-se selecionar de forma parcimoniosa apenas aquelas que possuem contribuição direta para determinação do desempenho a ser avaliado (Ferreira & Gomes, 2020).

A definição de métodos específicos para a etapa de seleção de variáveis não é amplamente desenvolvida na literatura (Senra et al., 2007), prevalecendo uma abordagem clássica de seleção de variáveis com base na opinião de especialistas, ou até mesmo na disponibilidade de dados. Tais abordagens, antecedentes à coleta de dados são, no entanto,

compensadas por uma série de recomendações metodológicas a partir do momento que já se dispõe dos dados, voltando-se para a exploração estatística e matemática da adequação das variáveis aos pressupostos e objetivos de análise (Senra et al., 2007). Tais métodos podem envolver diferentes graus de intervenção do pesquisador, uma vez que não se deve buscar um critério puramente matemático para definição do modelo, mas um que também seja coerente com o objetivo da análise. Dessa forma, para qualquer método utilizado, o pesquisador deve ser confrontado com o seu resultado e verificar a sua coerência com a realidade operacional e gerencial das DMUs (Senra et al., 2007).

Pode-se distinguir três procedimentos para a seleção de variáveis (Golany & Roll, 1989): 1) Seleção por Julgamento; 2) Análises estatísticas; e 3) Análises com modelos de DEA. A seguir, detalham-se, brevemente, algumas das possibilidades trazidas pela literatura em cada um desses estágios.

i. Seleção por Julgamento

Os primeiros aspectos a serem considerados para a definição das variáveis de um modelo de Análise Envoltória de Dados referem-se às características da atividade avaliada e ao objetivo da análise (Ferreira & Gomes, 2020). Um mesmo sistema organizacional pode ser avaliado sob diferentes perspectivas, ou dimensões, acerca das atividades realizadas e, sobretudo, acerca do resultado que ele proporciona ao ambiente (Senra et al., 2007). A literatura de avaliação educacional, por exemplo, avalia a eficiência de unidades de ensino sob variados pontos de vista, sendo os principais o educacional, tendo como referencial de desempenho a quantidade de estudantes formados (e.g. Giacomello & Oliveira, 2014), e o científico, enfocando a quantidade de publicações (e.g. Yang et al., 2018), podendo ser analisados de forma conjunta (e.g. Falquetto et al., 2018). Não obstante, há também perspectivas financeiras como a lucratividade, e psicossociais, que consideram a aprendizagem e a satisfação de alunos e efeitos amplos na sociedade e no mercado de trabalho (Johnes, Portela & Thanassoulis, 2017). Uma vez definidos com clareza a atividade avaliada e o objetivo da análise, os seguintes procedimentos clássicos são recomendados:

- 1) *Estudar detalhadamente o setor, ramo e atividade em análise;*
- 2) *Consultar especialistas do setor, ramo e atividade sobre quais são as variáveis (insumo e produtos) relevantes;*
- 3) *Utilizar um número restrito de variáveis, ou seja, apenas as mais relevantes.* (Ferreira & Gomes, 2020, pp. 142-143)

Sendo assim, o pesquisador deve buscar conhecer com profundidade a atividade a ser avaliada e buscar parâmetros para uma proposição preliminar de um modelo que ensinará a coleta de dados.

ii. Seleção por Análise estatística

O pesquisador, baseando-se nas etapas descritas acima, busca coletar os dados considerados relevantes para, então, analisá-los preliminarmente com ferramentas estatísticas. Nessa etapa, um procedimento fundamental perpassa pela análise da correlação entre as variáveis assumidas como relevantes (Ferreira & Gomes, 2020; Senra et. al., 2007). Tal etapa identifica variáveis com alta correlação entre si e enseja um exame acerca de quais dessas sejam, possivelmente, redundantes, na medida em que possuem “a mesma importância na determinação dos escores de eficiência técnica das DMUs” (Ferreira & Gomes, 2020, p. 143). Adicionalmente, além da análise de correlação, pode-se recorrer à análise de regressão para confirmação das relações entre insumos e produtos (Senra et. al. 2007), assim como abordagens mais avançadas como análises multivariadas, de componentes principais e análise fatorial (Ferreira & Gomes, 2020).

iii. Seleção por análises com modelos de DEA

A maior parte dos métodos de seleção de variáveis, no entanto, se concentrará no terceiro estágio de testagem iterativa de modelos para avaliar como as variáveis e seus resultados se comportam sob diferentes pressupostos de análise (Ferreira e Gomes, 2020). Essa análise constitui cânone da literatura de DEA, já descrito por Charnes et. al. (1978), no qual o pesquisador deve analisar a eficiência das DMUs sob a perspectiva de diferentes modelos de DEA, com diferentes combinações de variáveis, visando identificar a consistência dos resultados de eficiência sob diferentes pressupostos, assim como verificar quais variáveis são consistentemente relevantes para os resultados em diferentes modelos. Os modelos testados devem partir de diferentes pressupostos e modelagens em que, por exemplo, deve-se tornar algumas variáveis não-discrecionárias, ou seja, não ajustáveis pelo modelo de otimização, com intuito de verificar seu impacto nos resultados. Por fim, Senra et. al. (2007) variou métodos de avaliação multicritério para seleção de variáveis pela iteração de análises de eficiência.

Por fim, a seleção de variáveis também deve levar em consideração a definição do nível de complexidade que se deseja avaliar. Os modelos mais avançados de DEA permitem a definição de análises em rede ou a avaliação projetada no tempo. Nesse sentido, deve-se observar também quais os relacionamentos entre as variáveis de interesse e como elas se encadeiam, para que se tenha uma adequada representação do processo produtivo. Tais decisões

podem envolver a testagem de diferentes modelos, de forma que a maior parte das medidas para definição do modelo de Análise Envoltória de Dados de dará posteriormente à coleta com utilização dos dados disponíveis (Ferreira & Gomes, 2020; Senra et al., 2007)

O que se conclui, portanto, dos procedimentos para seleção e definição de variáveis para aplicação do DEA é que o conhecimento teórico e prático acerca da atividade analisada é fundamental, uma vez que, independentemente do método utilizado, o pesquisador deve ser capaz de avaliar a coerência das variáveis selecionadas e dos resultados obtidos com a realidade operacional e gerencial das DMUs investigadas (Senra et al., 2007).

Para o caso de campos consolidados, como é o caso da Educação no qual o DEA se originou e desde então possui extensiva aplicação (Costa et. al., 2012) é possível, a partir da análise da literatura e sobretudo por meio do estudo de revisões sistemáticas do campo, acessar as principais variáveis utilizadas como uma forma de antecipar variáveis que se mostraram relevantes em outros estudos. Dessa forma, tendo em vista o esforço do presente trabalho de desenvolver um modelo de avaliação da eficiência de unidades de treinamento, se recorrerá à literatura de Educação para compreensão das modelagens utilizadas, assim como a uma revisão da literatura de TD&E para compreensão das principais variáveis operacionalizadas nesse campo.

2.4 Avaliação da eficiência em Educação

Parte significativa das metodologias de avaliação formal tiveram sua origem no campo da Educação, remetendo a estudos sobre a efetividade de métodos de ensino e aprendizagem no início do século XX (Stufflebeam & Coryn, 2014). Variadas correntes e metodologias de avaliação surgem em diferentes contextos adequadas às diferenças institucionais e culturais entre países (Johnes, Portela & Thanassoulis, 2017; Sanders, 2003).

No que concerne à avaliação de programas educacionais, Sanders (2003) propõe um conceito integrador em que programas são compreendidos como “atividades contínuas de longo prazo que representam esforços coordenados e planejados para alcance dos principais objetivos educacionais” (Sanders, 2003, p. 697, tradução livre), propondo como exemplo, um programa de educação em artes de uma determinada escola. Tal definição se aproxima daquela utilizada para o conceito de Treinamento, com um esforço planejado e formal para capacitação de empregados para o alcance dos objetivos organizacionais, com a diferença de que esses são empreendidos por uma organização visando a melhoria do desempenho de objetivos relacionados ao contexto de trabalho (Goldstein, 1980).

Possivelmente, o que seria compreendido em TD&E como um programa de treinamento seja, para o campo da Educação, um conceito distinto. Sanders (2003) propõe o que seriam projetos educacionais, que se distinguem dos programas por serem “atividades de curto prazo com objetivos e recursos alocados específicos” citando, nesse caso, como exemplo o treinamento e desenvolvimento profissional do corpo docente de uma escola. O autor ressalta, no entanto, que o método de avaliação para ambos os casos, programas e projetos, é o mesmo, sendo utilizado, na mesma obra (Kellaghan, Stufflebeam & Wingate, 2003) o *framework* proposto por Stufflebeam (2003), o CIPP. Dessa forma, resta claro que os métodos de avaliação educacional, seja para programas ou projetos educacionais, se aplicam a esforços de Treinamento – como visto no caso do CIPP aplicado a ambos os campos – independente das possíveis diferenças conceituais relativas à noção de programas ou projetos que possa haver entre ambos os campos.

Uma outra similaridade entre Educação e TD&E reside na argumentação dos efeitos das ações educacionais. Conforme Aguinis & Kraiger (2009), os treinamentos podem legar efeitos para os indivíduos, equipes, organizações e para a sociedade como um todo. Tal asserção também é apresentada por Johnes, Portela & Thanassoulis (2017) e Grosskopf, Hayes & Taylor (2014a), em que evidências demonstram que a educação propicia o desenvolvimento econômico em nível nacional, desenvolvimento do capital humano e inovação nas organizações, assim como proporciona ganhos para os indivíduos, em que maior tempo de escolarização se reflete em ganhos crescentes por toda a vida. Johnes et al. (2017) ressaltam, no entanto, que estudos demonstram que a qualidade da educação é mais relevante para seus resultados do que a sua quantidade. No entanto, a operacionalização da qualidade dos estudos, normalmente abrangida pelo desempenho dos estudantes em testes padronizados, pode ser de difícil mensuração quando considerados contextos distintos, por exemplo, a comparação da qualidade da educação em diferentes escolas e, sobretudo, em diferentes regiões ou países (Johnes et al., 2017).

Conforme argumentam estudos acerca desse assunto, a maioria das avaliações educacionais, de forma similar ao TD&E, se volta primordialmente para seus efeitos, ou sua efetividade, e não para sua eficiência (Johnes et al., 2017; Lindsay, 1982). Johnes et al. (2017) definem para o campo educacional, o conceito básico de eficiência em que essa ocorre quando “os resultados observados da educação – como resultados de testes ou de valor agregado – são produzidos no nível mais baixo de recursos” (Johnes et al., 2017, p. 331, tradução livre). O conceito de eficiência trazido por Johnes et al. (2017), publicado em um periódico do campo de Pesquisa Operacional, se orienta para o conceito de eficiência técnica de Farrel (1957) e é

consoante ao conceito utilizado por outros estudos no campo educacional como os de Lindsay (1982), De Witte & López-Torres (2017) e Grosskopf, Hayes & Tayler (2014b), estando alinhado também ao conceito de eficiência dos campos da Avaliação (Suchman, 1967), Economia (Rosano-Peña, 2008) e da Administração Pública (Manzoor, 2018) e finalmente em TD&E (Abbad et al., 2020), ou seja, todos compreendem a eficiência em termos de melhor uso de recursos para alcance do máximo de objetivos ou resultados.

No entanto, no campo da Educação há diferentes abordagens para a mensuração da eficiência educacional que poderão enfatizar, por um lado, a aprendizagem, e por outro a gestão de recursos alocados – humanos, financeiros e humanos. Segundo Johnes et al. (2017), a eficiência pode se referir à melhor utilização recursos financeiros, mas também, por exemplo, à habilidade inata dos alunos (características da clientela). O entremeamento dessas dimensões também se observa no campo de TD&E, em que a maior parte dos modelos de avaliação, por exemplo o MAIS de Borges-Andrade (2006), abrange tantos insumos para a aprendizagem quanto recursos para a implementação do TD&E. A revisão de literatura de Renner, Laumer & Weitzel (2014), por exemplo, encontra apenas estudos que analisam a eficiência do TD&E em educação corporativa sob a perspectiva da aprendizagem (*learning efficiency*) conceituada como “o resultado da aprendizagem ou o ganho de conhecimento em relação ao tempo de aprendizagem” (Renner et al., p. 3), sendo esta analisada em função das políticas de treinamento e características instrucionais dos cursos corporativos.

Nesse sentido, torna-se relevante uma distinção entre eficiência de aprendizagem, em que insumos de natureza instrucional são enfatizados, e aquilo que nesse estudo denomina-se como a eficiência em perspectiva gerencial, em que os recursos financeiros, humanos e materiais são enfatizados. Essa distinção não impede a avaliação conjunta de insumos de ambas as naturezas, mas é útil para delineamento dos objetivos de avaliação e dos tipos de desempenhos – instrucional ou gerencial – que podem ser avaliados e, conseqüentemente, a natureza dos ajustes das ineficiências. Com isso, a perspectiva gerencial da eficiência, não tem como objetivo principal a avaliação da efetividade, remetendo, novamente, à função distinta e complementar da avaliação da eficiência (Suchman, 1967).

A diferença entre essas perspectivas em Educação é discutida, por exemplo, por Lindsay (1982) ao destacar que a avaliação do desempenho institucional de universidades deveria considerar ambas as dimensões, de efetividade e eficiência, sendo esta última relevante na medida em que recursos públicos para a educação tornam-se mais escassos (Sanders, 2003). Segundo o autor, a perspectiva da eficiência tende a ser negligenciada, haja vista que as usuais “autoavaliações” institucionais, comumente utilizadas nesse contexto, atestam antes pela

eficácia educacional, mas não seriam suficientes para atestar, sobretudo, a eficiência dos investimentos educacionais em perspectiva gerencial. Lindsay (1982) propôs, à época, que se buscassem abordagens quantitativas de insumo-produto baseados nas emergentes teorias econômicas, mas que essas deveriam considerar a multiplicidade de insumos e produtos das instituições educacionais e suas características intangíveis. Uma das proposições de Lindsay (1982) foi a perspectiva administrativa (*Administrator's Approach*) com a consideração de recursos financeiros, humanos e materiais, como insumos, e a quantidade de estudantes formados como produto, tendo também por consideração medidas sobre características institucionais e ambientais. Para considerar aspectos de mérito e valor associados à efetividade – ou qualidade do ensino, conforme Johnes et al. (2017) – Lindsay (1982) revisou variáveis adicionais para definição dos produtos que poderiam envolver os resultados individuais, como aprendizagem – e.g. notas em disciplinas – e satisfação ou mudança de atitudes – e.g. reação. O autor também distingue níveis de avaliação, que podem se deter sobre o aluno, o departamento ou a instituição como um todo, assim como diferentes tipos de resultados, os quais podem referir-se atividade de ensino-aprendizagem ou à produção científica.

A abordagem insumo-produto proposta por Lindsay (1982) foi abrangida pela incorporação de métodos de avaliação da eficiência, tais como o da Análise Envoltória de Dados – DEA desenvolvido pouco antes por Charnes et al. (1978) e que, segundo Johnes et al. (2017), promoveu um grande aumento da literatura sobre avaliação da eficiência educacional, tornando-se um dos métodos mais utilizados no campo. Desde então, conforme recentes revisões de De Witte & López-Torres (2017) e Suescún, Cubillos & García (2017), a avaliação da eficiência educacional baseada na utilização do DEA se tornou uma literatura ampla e desenvolvida com análises em todos os níveis e tipos de unidades educacionais, envolvendo: alunos, professores, escolas primárias, secundárias, universidades, distritos, estados nacionais e países (De Witte & López-Torres, 2017; Johnes et al., 2014). Atualmente, conforme De Witte & López-Torres (2017), no campo da Educação as literaturas de eficiência e efetividade são bastante distintas, sendo a eficiência educacional estudada também por um campo denominado de economia da educação (Johnes et al., 2017).

Para os níveis institucionais, as avaliações de eficiência apoiam a otimização da utilização de recursos para concretização do ideal de uma educação de qualidade com ampla cobertura da população (Suescún et al., 2017). Esta última, também referida como “equidade” (Grosskopf et al., 2014b; Johnes et al., 2017), consiste em um critério de grande interesse para a gestão de políticas públicas educacionais, complementar à avaliação da efetividade educacional.

Com efeito, no que concerne a políticas públicas, a literatura de finanças públicas utiliza as metodologias de Análise Envoltória de Dados para realização de grande quantidade de estudos acerca da qualidade do gasto público (Afonso, 2007; Boueri et al., 2007; Siqueira et al., 2020), sendo a Educação o campo mais estudado, conforme revisão de Sant'Anna et al. (2020), com destaque para a elevada produção científica de pesquisadores brasileiros. Groskopf et al. (2014a; 2014b) destacam – assim como Modesto (2000) e Afonso (2007) para o contexto nacional – a relevância da avaliação de eficiência para controle fiscal das atividades públicas que lidam com a falta de incentivos de mercado para se comportarem de forma eficiente e que são particularmente vulneráveis a comportamentos rentistas e disfunção burocrática (Groskopf et al., 2014b).

Com relação às políticas educacionais em específico, Groskopf et al. (2014a; 2014b), destacam o papel das avaliações de eficiência na orientação dos formuladores de políticas públicas que, muitas vezes, tomam decisões de alocação e gestão educacional sem apoio em teorias científicas da Educação e da Economia. Sob o ponto de vista educacional, as decisões de alocação de recursos, por exemplo, podem não levar em consideração os contextos econômicos, sociais e familiares dos alunos, os aspectos regionais sociodemográficos como nível de urbanização, disponibilidade de transportes e nível de renda, tampouco os níveis de formação e competência requeridos de professores e gestores educacionais. Sob o ponto de vista econômico e gerencial, devem ser testadas alternativas de provisão pública ou privada da educação, níveis ótimos de centralização ou descentralização, assim como a expectativa de que mecanismos de competitividade entre escolas promoverão maior desempenho (Groskopf et al., 2014a; 2014b). Os autores citam diversos estudos em eficiência com utilização do DEA que visam testar essas pressuposições para assegurar a eficácia das políticas educacionais e melhor retorno sobre os recursos dispendidos.

Dessa forma, a avaliação da eficiência educacional com a utilização do DEA pode almejar uma grande variedade de objetivos de avaliação, em diferentes níveis e tipos de resultados, tendo por consideração múltiplas variáveis de insumos, produtos e fatores contextuais, visando apoiar decisões em perspectiva educacional e gerencial, de educadores, organizações e governos, para o alcance da eficácia e a efetividade de atividades educacionais (Groskopf et al., 2014a; 2014b; Johnes et al., 2017). Visando sistematizar e analisar como a literatura de eficiência educacional tem abrangido essa variedade de elementos De Witte & López-Torres (2017) realizaram uma revisão de literatura com base na análise de 225 artigos científicos para descrição das principais variáveis de insumo, produto e contexto

operacionalizadas nesse campo, por meio de métodos paramétricos e não-paramétricos de avaliação da eficiência técnica de Farrel (1957).

A revisão de literatura sistematiza as principais variáveis para insumos, produtos e fatores contextuais, chamadas de determinantes da eficiência educacional e os efeitos observados que essas variáveis possuem na eficiência – positivos, negativos ou resultados mistos. Os autores apresentam primeiramente os “insumos discricionários especificados na função-produção da educação (ou aqueles fatores que são passíveis de gestão ao controle)” (De Witte & López-Torres, 2017, p. 3), divididos em quatro categorias: variáveis relacionadas ao estudante, à família, à instituição educacional e à comunidade. As mesmas categorias são utilizadas para classificação das variáveis exógenas, as “variáveis não-discricionárias ou ambientais que são responsáveis por diferenças nos resultados acadêmicos” (De Witte & López-Torres, 2017, p. 8), correspondendo àquelas que não são gerenciáveis, sendo operacionalizadas nos modelos matemáticos como variáveis fixas, não sujeita a ajustes e que funcionam como contingências à função-objetivo.

Os produtos são apresentados em categorização distinta, relacionadas aos tipos de resultados: 1) desempenho do estudante, 2) desempenho de atividade educacional e científica; 3) resultados educacionais; e 4) empregabilidade e desempenho de estudantes no mercado. O autor nota que as variáveis utilizadas na operacionalização dos produtos são relativamente similares entre os estudos, em que todos utilizam alguma medida de número de graduados, taxas de aprovação e as pontuações médias em testes. O número de graduados (80 artigos) captura a quantidade de produção educacional, mas não captura a qualidade, melhor refletida nas pontuações dos testes (127 artigos) (De Witte & López-Torres, 2017). A Tabela 4 apresenta, portanto, primeiramente as principais variáveis para insumos divididas nos respectivos níveis de análise.

Tabela 4
Variáveis de insumos utilizadas para avaliação da eficiência educacional

Estudante: Variáveis de insumo.
<p>1. Variáveis psicológicas e comportamentais</p> <ul style="list-style-type: none"> • Motivação e aspirações • Grupos de colegas (<i>Peer group</i>) • Desempenho previsto • Desempenho acadêmico prévio <p>2. Variáveis demográficas</p> <ul style="list-style-type: none"> • Deficiências (necessidades especiais de educação) • Pagamento por refeições escolares • Bolsas e auxílios • Idade, Gênero, Estado Civil • Idioma nativo (domínio limitado do Inglês) • Raça, Etnia, Minoria, Nacionalidade • Estilo de Vida
Família: Variáveis de insumo
<ul style="list-style-type: none"> • Necessidades econômicas • Estrutura familiar • Nível educacional dos pais • Relacionamento com crianças • Recursos disponíveis em casa, uso de <i>internet</i> • Status socioeconômico (Renda familiar, situação de emprego)
Instituição Educacional: Variáveis de insumo
<ul style="list-style-type: none"> • Taxa de aceitação (seleção de alunos) • Frequência • Clima escolar (<i>climate</i>) [similar a clima organizacional] • Índice de desistência • Recursos educacionais (livros, área construída, computadores, aulas, ônibus, bolsas). • Matrículas • Despesas (ensino, pesquisa, administração, equipe de apoio) • Proporção professor/aluno, número de professores, professores com doutorado. • Satisfação no Trabalho • Índice de mobilidade • Tipo de Organização: pública, privada, autônoma (<i>charter</i>). • Índice de visita parental • Pessoal (professores, equipe acadêmica, equipe de gestão, equipe de apoio, outros empregados) calculados em tempo de trabalho (<i>Full Time Equivalent – FTE</i>) • Receita de pesquisa, receita educacional (<i>tuition fees</i>), financiamento externo. • Tamanho (número de alunos, aluno por classe, proporção de meninos e meninas) • Proporção aluno/professor (ou vice-versa) • Absenteísmo de professores • Idade, gênero, raça de professores • Experiência ou formação de professores • Métodos de ensino, organização e gestão, qualidade e inovação de professores • Salário de professores
Comunidade: Variáveis de insumo
<ul style="list-style-type: none"> • Competição: índices de competitividade, número de instituições de ensino, localização • Características da vizinhança (impostos pagos, nível de empregos) • Porcentagem de domicílios com crianças em idade escolar • Porcentagem da população com educação pós-primária • Área urbana / rural (localização)

Nota. Fonte: De Witte & López-Torres (2017)

Como pode ser observado, os insumos cobrem em grande parte, características e variáveis de contexto relativas aos alunos e aos professores, envolvendo dados demográficos, educacionais, cognitivos e socioeconômicos. Há uma grande semelhança entre os níveis de análise de alunos e famílias com os insumos previstos em TD&E relativos, sobretudo, às características da clientela. Conforme Borges-Andrade et al. (2012, p. 24), os insumos para a avaliação de TD&E conforme o modelo MAIS, correspondem aos “fatores físicos e sociais e aos estados comportamentais e cognitivos, anteriores à instrução, que podem afeta-la ou os seus resultados”. O fato do modelo MAIS ter sido desenvolvido primeiramente para o contexto educacional (Borges-Andrade et al., 2012) pode ser uma das razões para essa compatibilidade. Segundo De Witte & López-Torres (2017), desempenho progresso, raça/etnia/minoria/nacionalidade são as variáveis mais utilizadas no nível do indivíduo e para o nível das famílias são mais utilizados como insumos o status socioeconômico e a educação dos pais, muito embora, todas as variáveis mencionadas tenham se mostrado como fortes preditoras do sucesso de estudantes.

O nível de análise das instituições educacionais é o mais numeroso em quantidade de estudos e de variáveis utilizadas. A análise de desempenho institucional se mostra também relevante em outras revisões de literatura específicas para esse nível, como a de Suescún et al. (2017) e Casado (2007). Os insumos nesse nível de análise sistematizados por De Witte & López-Torres (2017) capturam variados aspectos educacionais, gerenciais, estruturais e dados de pessoal e de professores e que são, normalmente, relacionados com resultados educacionais. As variáveis mais operacionalizadas são aquelas relativas aos gastos da instituição e variáveis de pessoal. Suescún et al. (2017), abrangendo recursos de pessoal, financeiros e materiais, também encontram prevalência de insumos relativos à quantidade de professores e de alunos na dimensão de pessoal, gastos de funcionamento e gastos com pessoal na dimensão financeira, e dimensão de recursos materiais, a área construída e os recursos instrucionais (livros, equipamentos).

Seguidamente aos insumos categorizados, De Witte & López-Torres (2017) apresentam os principais produtos, divididos conforme tipos de resultados. A Tabela 5 apresenta as principais variáveis para produtos sistematizadas pelos autores.

Tabela 5
Variáveis de produtos utilizadas para avaliação da eficiência educacional

Resultados dos Estudantes
<ul style="list-style-type: none"> • Número de formados (porcentagem de aprovados) • Desempenho dos alunos, Notas de dos alunos em diferentes disciplinas (leitura, línguas, matemática, artes)
Atividades de publicação e pesquisa
<ul style="list-style-type: none"> • Citações (impacto de pesquisa) • Contratos, patentes, prêmios e transferências tecnológicas • Créditos em disciplinas fornecidos pelo departamento • Teses de doutorado • Publicações: artigos publicados em revistas internacionais, livros, capítulos, ou outros produtos de pesquisa. • Outras atividades de pesquisa e ensino. • Qualidade do ensino e da pesquisa (classificação/índice/padrão de qualidade) • Bolsas de pesquisa recebidas, receita de pesquisa.
Resultados Educacionais
<ul style="list-style-type: none"> • Frequência • Índice de desistência • Matrículas • Refeições servidas, número de leitões. • Pessoal ou alunos em atividade em outros países (<i>overseas staff/student</i>) • Atitudes de professores • Receitas educacionais
Resultados no Mercado
<ul style="list-style-type: none"> • Empregabilidade de estudantes (graduados com empregos, local de emprego) • Salário inicial de graduados • Satisfação de estudantes (questionários)

Nota. Fonte: De Witte & López-Torres (2017)

Conforme mencionado, as variáveis de produtos são relativamente homogêneas e se concentram em resultados para os estudantes, em termos de número de formados, aprendizagem ou desempenho. Uma quantidade menor de estudos inclui também a satisfação ou sucesso no mercado de trabalho de estudantes. Também é frequente a consideração de variáveis para produtos de pesquisa em variados estudos.

Muitas das variáveis educacionais são mais utilizadas como insumos (Tabela 4), de forma que são poucos os estudos que consideram variáveis de resultados classificados como educacionais na Tabela 5 e não há registro de resultados para o nível da comunidade. De Witte & López-Torres (2017) argumentam que o campo termina por focar resultados de curto e médio prazo, não se buscando avaliar a eficiência em relação a resultados de longo prazo. Os resultados de De Witte & López-Torres (2017) são consistentes com os resultados de Suescún et al. (2017) que identifica como principais produtos na literatura revisada: a quantidade de

alunos formados, quantidade de publicações científicas, notas de estudantes em testes padronizados de governo (*Pruebas del Estado*), horas de ensino e créditos ofertados e receitas de projetos de pesquisa e desenvolvimento. Como pode-se verificar, o critério principal investigado pela literatura de eficiência educacional são os resultados no nível dos indivíduos (produtos).

O enfoque no nível individual do campo de avaliação educacional observado por De Witte & López-Torres (2017) guarda similaridade com o escopo usual também no campo de TD&E em que estudos de avaliação nos níveis de resultados individuais – reação, aprendizagem e desempenho – possuem relevante produção na literatura (Alvarez et al., 2004; Cassiano & Borges-Andrade, 2017; Sampaio et al., 2018; Kraiger & Ford, 2021) e são mais utilizadas nas organizações (Kennedy et al., 2014). A partir dos modelos integrativos, o campo de TD&E também delineou variáveis contextuais que afetam a efetividade do Treinamento, de forma que há também nesse sentido uma similaridade teórico-metodológica com o campo educacional. Para a literatura de avaliação da eficiência educacional, essa perspectiva se consubstancia na análise de fatores exógenos ou variáveis não-discrecionárias.

Conforme mencionado anteriormente, a análise de variáveis exógenas consiste no método de dois estágios (Afonso, 2007; Boueri et al., 2007), no qual após o cálculo da eficiência (primeiro estágio) são utilizadas técnicas estatísticas complementares para determinação da significância e sentido da influência (positiva ou negativa) de variáveis do ambiente ou não controladas pelos tomadores de decisão nos resultados de eficiência. As variáveis não-discrecionárias encontradas na literatura de avaliação da eficiência educacional são apresentadas por De Witte & López-Torres (2017) acompanhadas do efeito geral observado sobre os escores de eficiência técnica, sendo também chamados de condicionantes (Suescún et al., 2017). Para Cordero, Santín & Simancas (2017), essa abordagem pode ser chamada de avaliação da eficiência condicionada. As principais variáveis não-discrecionárias sistematizadas por De Witte & López-Torres (2017) com base nos estudos que aplicaram esse método são apresentadas na Tabela 6 e divididas nas categorias propostas pelos autores.

Tabela 6

Variáveis não-discrecionárias utilizadas para avaliação da eficiência educacional

Variáveis não-discrecionárias por nível de análise	Efeito sobre a Eficiência Educacional
Estudantes	
• Deficiências (necessidades especiais de educação)	Se for maior, menor o desempenho do aluno (e maiores os custos.)
• Pagamento por refeições escolares	Resultados mistos: positivos e negativos
• Gênero	Resultados mistos: positivos e negativos
• Bolsas	Resultados mistos: positivos e negativos
• Idioma nativo (domínio limitado do Inglês)	Menor domínio do inglês reduz desempenho
• Desempenho acadêmico prévio	Quanto maior, melhor os resultados dos alunos
• Raça, Etnia, Minoria, Nacionalidade	Se for imigrante, menor o desempenho do aluno.
Família	
• Estrutura familiar	Resultados mistos: positivos e negativos
• Nível educacional dos pais	Se for menor, piores os resultados do aluno.
• Relacionamento com crianças, envolvimento em casa.	Quanto maior, melhores os resultados do aluno.
• Recursos disponíveis em casa, uso de <i>internet</i>	Se menor, piores os resultados do aluno.
• Status socioeconômico (Renda familiar, situação de emprego)	Se menor, piores os resultados do aluno.
Instituição Educacional	
• Frequência	Se menor, menores os resultados dos alunos.
• Práticas de contratação	Práticas descentralizadas de contratação são melhores.
• Financiamento local ou externo (receitas de mensalidades de estudantes)	Quanto maior, melhor os resultados de eficiência.
• Tipo de Organização: pública, privada, autônoma (charter).	Resultados mistos: positivos e negativos
• Qualidade do ensino, pesquisa e inovação.	Quanto maior, melhores os resultados dos estudantes.
• Taxas de expulsão e suspensão de alunos.	Quanto maior, piores os resultados dos estudantes.
• Clima ou Orientação Religiosa da Instituição	Não é significativo em todos os casos.
• Tamanho (número de alunos, tamanho das turmas, proporção aluno/professor)	Resultados mistos: grandes instituições podem reduzir custos, mas seus resultados são piorados.
• Estrutura (matrícula / proporção de meninos e garotas)	Resultados mistos: positivos e negativos.
• Características do professor (idade / gênero / escolaridade / experiência / número / salário)	Resultados mistos: positivos e negativos. Não é significativo em todos os casos.
Comunidade	
• Competição: índices de competitividade, número de instituições de ensino dentro de X quilômetros.	Resultados mistos: positivos e negativos.
• Produto Interno Bruto per capita.	Quanto maior, melhores os resultados.
• Proporção de imigrantes	Quanto maior a proporção, piores os resultados.
• Taxa de mortalidade, crime e violência	Resultados mistos: positivos e negativos. Não é significativo em todos os casos.
• Características da vizinhança (oportunidades de emprego, acesso a riqueza, níveis de pobreza)	Se for prejudicial, piores os resultados dos alunos. Se houver escassez de empregos, piores os resultados dos alunos
• Porcentagem de domicílios com crianças em idade escolar	Resultados mistos: positivos e negativos.
• Porcentagem da população com / sem ensino superior	Quanto maior, melhores os resultados dos estudantes.
• População, tamanho do distrito educacional	Resultados mistos: positivos e negativos.
• Área urbana / rural (localização)	Instituições educacionais urbanas obtêm melhores resultados e podem reduzir custos

Nota. Fonte: De Witte & López-Torres (2017)

Comparando-se com a Tabela 4, relativa aos insumos, pode-se verificar, como também observado por De Witte & López-Torres (2017), que algumas variáveis de insumos são tratadas

como variáveis não-discrecionárias, sendo essa uma definição de modelo que dependerá do objetivo da avaliação.

As variáveis apresentadas, portanto, são utilizadas para análises em diferentes níveis e podem abranger perspectivas de desempenho de aprendizagem, econômico e gerencial (Johnes et al., 2017; Lindsay, 1982). Os resultados de De Witte & López-Torres (2017) sugerem que a perspectiva da eficiência educacional, ancorada em resultados de aprendizagem e foco no nível individual é prevalente na literatura. Conforme o autor, apenas 09 estudos em nível de sistema educacional foram encontrados. Possivelmente, conforme comentado anteriormente, tais perspectivas mais amplas sejam mais abrangidas por outras literaturas como os campos da economia, finanças públicas e de avaliação de políticas públicas (Afonso, 2007; Boueri et al., 2007; Jannuzzi, 2015) que se voltam para a avaliações com os objetivos apontados por Groskopf et al (2014a; 2014b) de subsidiar cientificamente as decisões governamentais de financiamento e gestão de políticas educacionais.

De forma complementar, portanto, ao estudo de De Witte & López-Torres (2017), a revisão de Siqueira et al. (2020) sobre as principais variáveis de insumos e produtos operacionalizados em avaliações da eficiência do gasto público em diferentes funções de governo, destaca aquelas utilizadas na avaliação da Educação (Tabela 7).

Tabela 7

Relação de insumos e produtos utilizados em eficiência na Educação.

Insumos	Produtos
<ul style="list-style-type: none"> •Gastos com educação •Gastos com educação <i>per capita</i> •Gastos com funcionários •Gastos com investimentos •Proporção aluno/professor •Índice de qualidade da educação •Alunos matriculados 	<ul style="list-style-type: none"> •Áreas de espaços recreacionais •Índice de aprovação •Índices e sub-índices setoriais •Notas em índices de avaliação da educação (IDEB, PISA, etc). •Número de matrículas •Notas no idioma nativo e matemática •Número de escolas •Número de pessoas com 4 anos de estudo •Número ou porcentagem de pessoas atendidas ou inscritas em programas sociais. •Percentual de alfabetizados em determinada faixa etária. •Taxa de crescimento do PIB.

Nota. Fonte: Siqueira et al. (2020)

Verifica-se que os gastos educacionais, estratificados em diferentes categorias, naturalmente, são a principal categoria de insumos, os quais consideram, no entanto, alguns aspectos observados no nível da instituição educacional como o fluxo de entrada de alunos, a

proporção aluno/professor e índices de qualidade da educação, oriundos de outros tipos de avaliações que são combinadas com a avaliação da eficiência – como proposto por Casado (2007). Os produtos, por sua vez, incluem aspectos tradicionais, como o desempenho de alunos, vistos também em grande quantidade de estudos revisados por De Witte & López-Torres (2017). No entanto, avançam na inclusão de índices que se refeririam ao nível social mais amplo, não observado naquela revisão. Assim, Siqueira et al. (2020) apresentam uma literatura complementar que demonstra a utilização dos métodos de avaliação da eficiência orientados a benefícios citados por Suescún et al. (2017), Grosskopf et al. (2014b) e Johnes et al. (2017), tais como a melhoria de índices sociais por meio de políticas educacionais de qualidade com ampla cobertura da população (Suescún et al., 2017). O critério da equidade ou alcance da educação, por exemplo, encontra-se operacionalizado pelo quantitativo de escolas, pessoas atendidas ou que receberam níveis adequados de educação em determinada região. Dessa forma, tais pesquisas abrangem tanto o aspecto qualitativo, quanto o quantitativo da avaliação educacional (De Witte & López-Torres, 2017; Johnes et al., 2017), porém com objetivos de aperfeiçoamento menos orientados para a aprendizagem individual, mas para a gestão do sistema educacional (instituições) e das atividades governamentais (políticas educacionais).

Tendo sido apresentadas as variáveis mais relevantes da literatura de eficiência educacional, Johnes et al. (2017) apontam que revisões dessa natureza se constituem um recurso para pesquisadores desenvolverem modelos coerentes, buscando operacionalizar variáveis de forma similar e comparável com a literatura favorecendo a investigação de lacunas. Não obstante, segundo Suescún et al. (2017), a despeito da grande quantidade de variáveis, em sua revisão de 59 estudos, 52,54% deles utilizam apenas três ou quatro indicadores de insumo e 49% dois ou três indicadores de produtos, de forma que um modelo de análise nesse contexto utilizará, tipicamente, cerca de sete variáveis. Os autores também identificam que a maioria dos estudos (62%) realizam análises com orientação para o produto – ajustes para aumento de resultados – e em sua maior parte combinam os métodos CCR e BCC, ou seja, avaliam se as relações insumo-produto possuem retornos constantes (CRS – *Constant return to scale*) ou variáveis (VRS – *Variable return to scale*) à escala (Rosano-Peña, 2008).

Remetendo aos modelos avançados de DEA, uma das formas de se abranger mais variáveis, no entanto, é a especificação de um modelo em rede, discutido anteriormente, de forma a capturar mais aspectos da complexidade do sistema analisado (Kao 2017). As revisões de De Witte & López-Torres (2017) e Suescún et al. (2017) não deixam claro quantos e quais estudos utilizam esse tipo de modelagem. Para ilustrá-la, no entanto, dois estudos de avaliação da eficiência com *Network* DEA podem ser descritos.

O estudo de Falquetto et al. (2018), realizado no Brasil, por exemplo, utilizou um modelo de *Network* DEA em três estágios, com pressupostos de retornos constantes de escala (CRS) e orientação ao produto, para avaliar a eficiência produtiva de 34 programas de pós-graduação em economia beneficiados pelas políticas públicas do Programa de Excelência Acadêmica (Proex) e do Programa de Apoio à Pós-graduação (Proap). Conforme diagrama apresentado pelos autores (Figura 22), o modelo de análise considerava no primeiro estágio a eficiência na geração de matrículas (produtos intermediários) em relação ao número de professores e ao custeio (investimento) recebido da Capes (insumos). No segundo estágio verificou-se a eficiência na formação de mestres e doutores, assim como a publicação em anais de eventos científicos (produtos intermediários), com base no número de matriculados do estágio anterior. Por fim, o terceiro estágio analisou a eficiência quanto às publicações nos artigos e capítulos de livros ponderados pelo peso Capes (produtos finais). Os dados foram coletados em bases de governo e por meio de solicitações às universidades (Falquetto et al., 2018).

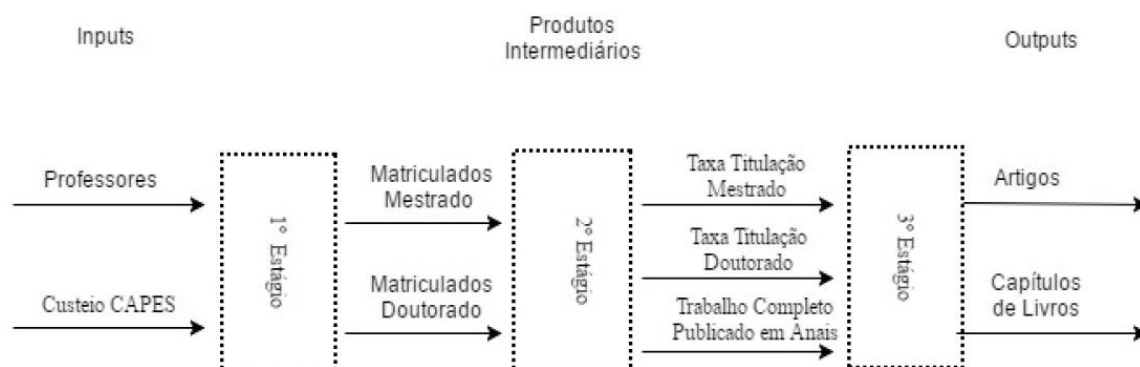


Figura 22. Modelo DEA para avaliação de programas de pós-graduação
Fonte: Falquetto et al. (2018)

Um segundo estudo, realizado por Yang, Fukuyama & Song (2018), utilizou um modelo *Network* DEA para determinar as ineficiências e a produtividade de 64 universidades de pesquisa. O estudo inovou ao adotar a perspectiva longitudinal, com dados relativos ao período de 2010-2013, para o qual desenvolveram uma medida para a evolução da produtividade (Yang et al., 2018) – uma alternativa ao índice de Malmquist (Ferreira & Gomes, 2020). O modelo proposto por Yang et al. (2018) concebe o processo de pesquisa como sendo composto de dois estágios (Figura 23). O primeiro estágio refere-se à eficiência das atividades de ensino e pesquisa, em que são insumos os fundos recebidos por entidades externas para que a universidade realize pesquisas e o financiamento público total provido principalmente para financiamento de atividades de ensino e manutenção da universidade. São considerados

produtos desse primeiro estágio a quantidade de alunos formados – graduados, mestres e doutores – e as publicações de alto impacto oriundas dessa atividade, assim como os registros de patentes e outras formas de propriedade intelectual (e.g. licenças de *softwares*) que servirão de insumo para o segundo estágio. O segundo estágio visa avaliar a eficiência da universidade em promover a inovação por meio de transferências tecnológicas, de forma que são considerados como insumos os registros de propriedade intelectual, mas também o quantitativo de pessoal alocado às equipes de serviços relacionados à transferência tecnológica. Por fim, os resultados da segunda etapa têm como produtos a receita oriunda da venda de patentes e transferências tecnológicas para outras organizações, em que uma parte – uma taxa definida pelo governo – é reinvestida, tornando-se insumo para a primeira etapa, sendo essa proporção definida em 50% (Yang et al., 2018).

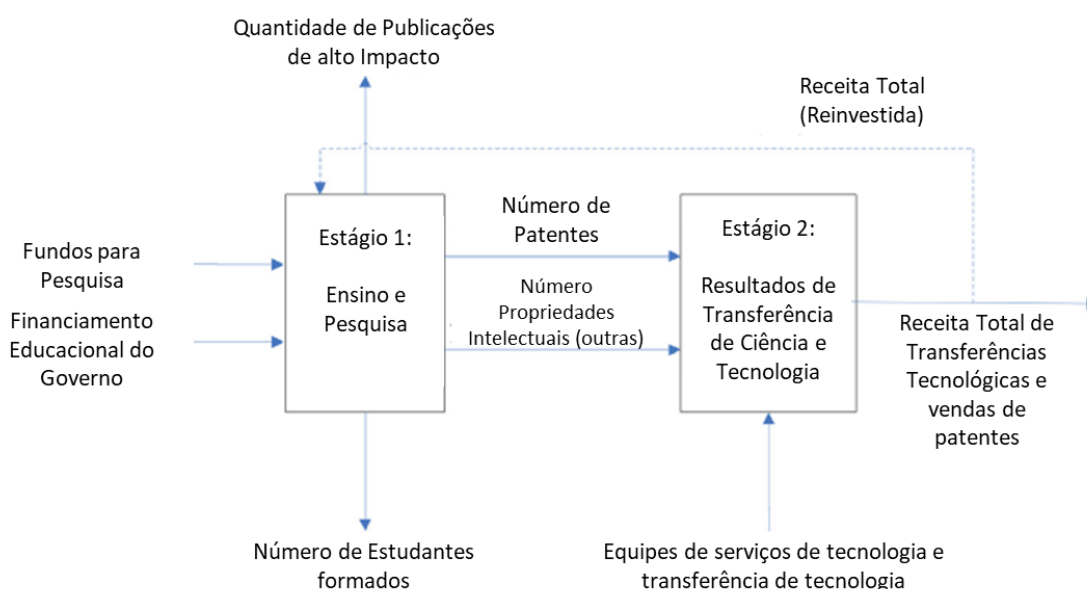


Figura 23. Modelo DEA para avaliação de universidades de pesquisa

Fonte: Yang et al. (2018), tradução livre.

Os dois artigos apresentados servem ao propósito, portanto, de ilustrar aplicações avançadas de modelos de Análise Envoltória de Dados em rede realizadas no campo da Educação, tendo como enfoque os níveis mais amplos das instituições e do sistema educacional, assim como a sua contribuição para decisões de investimento nessa atividade.

A seção 2.2 que versou sobre os modelos de avaliação em TD&E demonstrou como a concepção sistêmica do Treinamento permite a definição dos seus componentes em termos de um fluxo, nos quais se identificam suas principais variáveis de insumos e produto. Dentre os insumos, a dimensão dos recursos, embora abrangida por poucos modelos de avaliação da área,

serviu de fundamento para a proposição de um modelo teórico de avaliação de TD&E, que se alinha aos pressupostos das teorias de Avaliação (Suchman, 1967) e de como ela é aplicada em outros campos, sobretudo o da Educação (Stufflebeam, 2003). Uma diferença relevante, no entanto, reside nos resultados finais do Treinamento, que se voltam para as organizações. Não obstante, o campo da Educação, com ampla literatura consolidada (De Witte & Tórrez-Lopes, 2017; Johnes et al., 2017) pode servir de base para a aplicação da avaliação da eficiência do Treinamento em diferentes níveis: individual, equipes, organizacional e o nível mais amplo da sociedade.

Tendo em vista a proposição do modelo teórico de avaliação da eficiência do Treinamento e observados os resultados da literatura de avaliação educacional, o presente estudo avança em direção à avaliação, em perspectiva gerencial, da eficiência do treinamento abrangendo os níveis correlatos, que poderiam ser compreendidos como os relativos às unidades organizacionais responsáveis pela implementação de programas de treinamento para alcance de objetivos nas organizações nas quais estão inseridas. Para esse nível, também se observam, assim como na Educação, políticas públicas orientadas para o treinamento de trabalhadores para o mercado de trabalho (Mourão & Borges-Andrade, 2005) ou no âmbito das organizações públicas (OCDE, 2017), de forma que os investimentos realizados também podem ensejar avaliações de eficiência do gasto público como as observadas no campo da Educação (Johnes et al., 2017; Sant'Anna et al., 2020; Siqueira et al., 2020).

Dentre as variadas revisões de literatura apresentadas sobre avaliação da eficiência, no entanto, não foram encontradas revisões acerca da avaliação da eficiência em TD&E. Tendo em vista as possibilidades de desenvolvimento dessa perspectiva de avaliação para o campo, se buscou, ao lado das variáveis revisadas no campo da avaliação de eficiência educacional aqui apresentadas, identificar possíveis variáveis utilizadas em estudos empíricos de avaliação da eficiência em TD&E. Nesse sentido, a próxima seção se trata de uma revisão de literatura com o intuito de delinear variáveis, em perspectiva de insumo-produto para avaliação da eficiência em TD&E, visando subsidiar a operacionalização de um modelo DEA aplicado ao Treinamento.

2.5 Revisão de literatura: Eficiência das atividades de treinamento

A presente revisão de literatura buscou realizar uma pesquisa bibliográfica, tendo sido utilizada as etapas da metodologia *Methodi Ordinatio* proposta por Pagani, Kovalski & Resende (2015). Ao longo do desenvolvimento do trabalho, o reduzido número de artigos encontrados ensejou a supressão das etapas de cálculos dos índices de relevância de artigos – o índice InOrdinatio – com base em seu fator de impacto (FI), para que se considerasse todos os

artigos encontrados independente de seu FI. Como será visto adiante, poucos estudos se enquadraram nos critérios de pesquisa que buscavam pesquisas sobre eficiência de programas de treinamento em organizações sob uma perspectiva gerencial (recursos).

A revisão de literatura teve como objetivo geral levantar as principais variáveis descritivas para avaliação de programas de treinamento em contexto organizacional utilizadas em estudos empíricos nacionais e internacionais. Assim, como ponto de partida, a pesquisa iniciou-se com a definição preliminar do tema de interesse e a realização de uma pesquisa exploratória (Pagani et al., 2015).

A intenção de pesquisa traduziu-se no seguinte tema de pesquisa: “Indicadores de Eficiência de Treinamento” (em inglês: “*Training Efficiency Indicators*”), para levantamento de variáveis que apoiassem o objetivo de pesquisa. Poucos estudos possuíam como objetivo a avaliação da eficiência de treinamento, de forma que, para inclusão na revisão, bastava a descrição de quaisquer variáveis do processo ou programa de treinamento em nível organizacional. Como objetivo adicional, acrescentado posteriormente frente ao reduzido número de resultados nas pesquisas iniciais, buscou-se identificar estudos nos quais tenha sido utilizada a Análise Envoltória de Dados como método de avaliação de unidades ou programas de treinamento.

Não obstante, o critério para o recorte da pesquisa foi a consideração de estudos que se enquadrassem dentro da definição de “Treinamento” como um esforço planejado por uma organização para promover a aprendizagem de empregados de comportamentos relacionados ao trabalho com vistas à melhoria de desempenho individual e organizacional (Goldstein, 1980; Latham, 1988; Wexley, 1984). A definição é relevante uma vez que o “Treinamento” é um termo transversal a múltiplos campos de conhecimento (Garavan et al., 2019), inclusive para além das ciências organizacionais e da psicologia, sendo utilizado com diferentes significados e contextos de aplicação.

Com relação aos parâmetros de busca nas bases, definiu-se, conforme procedimentos recomendados de revisão de literatura (Pagani et al., 2015), que deveriam ser considerados apenas artigos científicos publicados em periódicos revisados por pares, visando o maior rigor e relevância dos resultados. Também se considerou que o objetivo da pesquisa visa identificar, prioritariamente, estudos empíricos (Campos, Pagani, Resende & Pontes, 2018). Para definição das bases de dados, buscou-se como referência outras revisões de literatura realizadas nas ciências organizacionais e na área de treinamento (e.g. Borges-Andrade & Pagotto, 2010; Gondim, et al., 2019; Pagani et al., 2015; Tasca et al., 2010). Dessa forma: foram definidas 10 bases para a realização das buscas: 1) Web of Science; 2) Science Direct; 3) Scopus; 4)

Proquest¹; 5) Emerald; 6) JSTOR²; 7) Sage Journals Online; 8) Wiley Online Library; 9) Scielo³; e 10) Spell⁴.

Com relação ao período de publicação, ao longo da pesquisa, optou-se por não se especificar um período (Pagani et al., 2015). Os campos de busca utilizados foram os de “Título”, “Resumo” e “Palavra-Chave”, adotando-se, quando necessária a escolha de um entre esses campos, aquele que retornasse mais resultados. Como explicitado por Pagani et al. (2015), as bases variam com relação às ferramentas de busca, devendo-se sempre manter a equivalência dos parâmetros, termos e combinações. Por exemplo, nem todas as bases permitem, ou realizam adequadamente, pesquisas com operadores *booleanos*. Sendo assim, as palavras-chave utilizadas consideravam suas variações e as ferramentas à disposição em cada base de dados⁵.

Para definição de palavras-chave, buscou-se sua derivação a partir da intenção de pesquisa (Pagani et al., 2015). Deve-se atentar para os idiomas da busca e para sinônimos e variações de cada termo. Para a presente pesquisa optou-se por uma busca ampla em três idiomas: inglês, português e espanhol. A partir da intenção de pesquisa definiu-se as seguintes palavras-chave: “Treinamento”, “Eficiência” e “Indicador” (“*Training*”, “*Efficiency*”, “*Indicator*”) e variações desta última por meio dos termos “Medida” e “Métrica” (“*Measure*”, “*Metric*”). A testagem desses termos na pesquisa exploratória retornou uma grande quantidade de artigos, mas poucos atinentes aos objetivos de pesquisa. Com isso, houve a ampliação do escopo da pesquisa para incluir possíveis artigos sobre Treinamento publicados no campo da Educação, uma vez que o termo “Educação Corporativa” (*Corporate Education*) estaria abrangido por essa busca. Por fim, considerou-se relevante buscar diretamente artigos que explicitassem como palavra-chave a metodologia de interesse, a Análise Envoltória de Dados (*Data Envelopment Analysis* – DEA). Dessa forma, em inglês, foram testadas 14 combinações (“AND”) de dois e três termos (Tabela 8).

¹ Bases do ProQuest utilizadas: Applied Social Sciences Index & Abstracts (ASSIA); ASFA: Aquatic Sciences and Fisheries Abstracts (1971 - atual); Coronavirus Research Database; ERIC (1966 - atual); Library & Information Science Abstracts (LISA) (1969 - atual); National Criminal Justice Reference Service (NCJRS) Abstracts Database (1975 - atual); PTSDpubs (1871 - atual); Research Library; SciTech Premium Collection (1946 - atual); Social Science Database; Sociological Abstracts (1952 - atual); Sports Medicine & Education Index (1970 - atual); Technology Collection (1962 - atual).

² Além da busca tradicional, a JSTOR permite a busca em bases dedicadas a tópicos de pesquisa. Dessa forma, realizou-se buscas nas páginas dos seguintes tópicos da JSTOR: “*Training*”, “*Educational Research*” e “*Efficiency Metrics*”. Os resultados foram computados em separado como uma nova base.

³ Busca foi realizada no portal acessado por meio do Periódicos CAPES ao invés do portal de acesso livre, devido à maior quantidade de ferramentas de busca e filtragem no primeiro.

⁴ Busca realizadas sem filtros e sem operadores devido à incongruência dos resultados, mas com inclusão de variações, tais como de gênero, número e flexões verbais, para as palavras-chave.

⁵ Science Direct e Emerald Insight possuem sistemas de busca sem operadores de *booleanos* de truncagem, porém os resultados incluem automaticamente variações dos termos utilizados, o que não prejudicou a busca.

Tabela 8

Termos de busca exploratória em bases de dados

Linha 1: Treinamento (T)	Linha 2: Educação (E)	Linha 3: DEA (D)
Training, Efficiency, Indicator Training, Efficiency, Measure Training, Efficiency, Metric Training, Efficiency Training, Indicator Training, Measure Training, Metric	Education, Efficiency, Indicator Education, Efficiency, Measure Education, Efficiency, Metric	Training, Data Envelopment Analysis Education, Data Envelopment Analysis Training, DEA Education, DEA

Nota. Fonte: Elaboração própria.

Para abrangência de sinônimos (e.g. “capacitação”) e variações, em inglês, português e espanhol foram utilizadas, afinal, 58 combinações distintas entre os termos de pesquisa. Após a etapa de busca exploratória, definiu-se as palavras-chave e bases de dados definitivas para a pesquisa (Pagani et al., 2015). Para o presente caso a pesquisa definitiva foi realizada com parte das combinações testadas na etapa exploratória, mantendo-se os parâmetros e mesmas bases de dados. Os resultados da pesquisa bibliográfica se referem ao período de maio e junho de 2020.

A pesquisa definitiva resultou em um total de 33.543 artigos em todas as dez bases de dados consultadas, ensejando os procedimentos de filtragem propostos por Pagani et al. (2015), que preveem o apoio de *softwares* de organização de referências como *Mendeley* e *JabRef*. Os procedimentos de filtragem consistem na análise individual de cada resultado de pesquisa e eliminação de artigos por meio dos seguintes passos e critérios: i) artigos duplicados; ii) trabalhos não científicos ou referentes a livros, capítulos e anais de conferências; e iii) artigos cujo título não se relaciona ao tema de pesquisa. Deve-se avaliar, também, os resumos e palavras-chave em casos em que o título não seja suficiente para avaliar a adequação do artigo ao escopo da pesquisa (Pagani et al., 2015).

Nessa etapa, optou-se pela reprodução de procedimento utilizado por Tasca et al. (2010), no qual palavras-chave eram utilizadas para buscar títulos de artigos com maior possibilidade de representar temas alheios ao interesse de pesquisa. Sendo assim, foram utilizadas cerca de 36 palavras-chave representativas de temas nas áreas da Educação Física, Esportes, Ciências Médicas, Engenharias, Agricultura, entre outros, na ferramenta de busca do *software Mendeley*, propiciando a exclusão mais rápida de grupos de artigos nessas temáticas. Além de artigos nesses campos, foram excluídos, também, artigos com temas educacionais em nível escolar e com enfoque em desempenhos de professores e alunos (e.g. *Pre-School*, *Secondary Education*, *Student Performance*, *Teaching*), assim como artigos sobre treinamentos em contextos não organizacionais. Por fim, excluiu-se uma grande quantidade de artigos em outros temas recorrentes, tais como pesquisas de avaliação econômica, avaliação de Índices de

Desenvolvimento Humano, *Machine Learning* (e.g. treinamento de inteligência artificial) e utilizações do DEA em campos distintos do interesse de pesquisa (e.g. desempenho do setor bancário, de aeroportos, redes elétricas e seleção de fornecedores). Ao final desse procedimento de filtragem, foram excluídos 9.737 artigos, reduzindo-se a amostra para 5.915 resultados.

Conforme sugerido por Pagani et al. (2015), em razão do alto número de artigos restantes, prosseguiu-se à leitura dos resumos de cada artigo para a identificação e exclusão daqueles com características distintas do escopo de pesquisa. Nessa etapa, além da identificação e exclusão de mais artigos com temas distintos da intenção de pesquisa, foram também identificados artigos teóricos e de revisão de literatura alinhados ao tema de interesse, mas que foram separados da amostra final, a qual visava priorizar artigos empíricos. Foram também separados da amostra os artigos que se referiam à análise com enfoque exclusivo em variáveis instrucionais e de aprendizagem no nível dos indivíduos.

Como resultado, foram excluídos mais 4.359 artigos, chegando-se a uma amostra de 1.556 artigos. Durante a etapa de leitura dos resumos procedeu-se à classificação desses artigos em relação ao tema principal da análise. Foi identificado que artigos retornados nas buscas sobre “Treinamento” e “Educação” frequentemente não tratavam de Treinamentos em contexto organizacional, como definido na intenção de pesquisa. Sendo assim, os artigos foram analisados e separados em sete categorias temáticas. Os artigos classificados como i) “Treinamento em Organizações” foram considerados para a revisão de literatura, enquanto as demais categorias foram excluídas, versando sobre ii) Gestão de Pessoas; iii) Capital Intelectual; iv) Gestão Organizacional; v) Gestão de Políticas Públicas; vi) Eficiência de Instituições de Ensino Superior (IES) e entidades de Pesquisa e Desenvolvimento - P&D; e vii) Treinamento Vocacional. A Figura 24 se propõe a ilustrar sumariamente esse processo e apresentar os quantitativos finais classificados em cada categoria temática emergente.

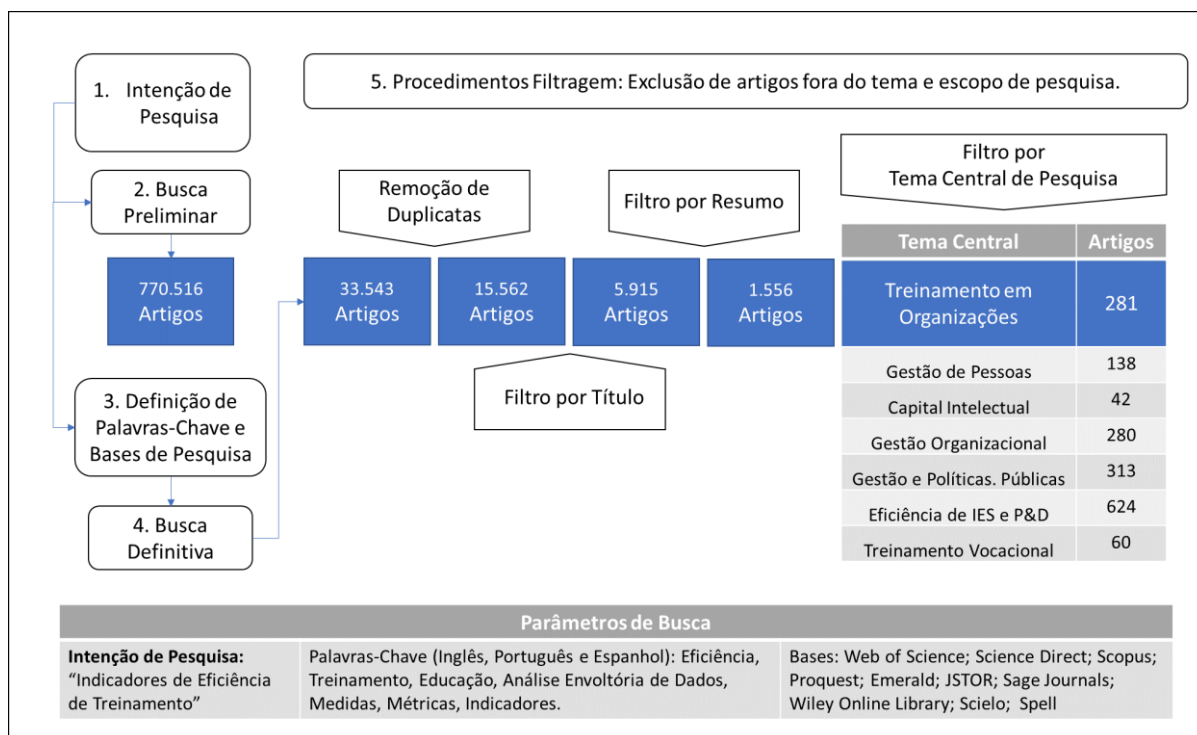


Figura 24. Sumário de procedimentos de busca e filtragem de artigos

Fonte: Elaboração própria.

Como pode-se verificar na Figura 24, a categoria “Treinamento em Organizações” é composta por 281 artigos, referindo-se àqueles estudos considerados como atinentes à intenção de pesquisa. Os artigos nessa categoria apresentavam pouca utilização do DEA como método de análise. Por esse motivo, adotou-se um critério menos rígido para inclusão na amostra de interesse de artigos que utilizassem a Análise Envoltória de Dados em contextos análogos ao de treinamento, como o de educação profissional, desde que operacionalizadas medidas relativas aos cursos fornecidos. Sendo assim, artigos que relacionavam a eficiência de Escolas de Negócios e programas de *Master in Business Administration* – MBA, por exemplo, foram incluídos como forma de se conhecer como se davam as aplicações do DEA nesse campo correlato. Não obstante, todos os 281 artigos foram baixados e lidos na íntegra.

A leitura dos artigos consistiu em novo processo de filtragem adotando-se os mesmos critérios das etapas anteriores, sobretudo, atinência ao tema e operacionalização de variáveis descritivas da atividade de treinamento. Após a leitura foram excluídos: 01 artigo não encontrado para *download*; 01 artigo de anais de evento; 03 artigos em línguas estrangeiras sem possibilidade de tradução segura (chinês, russo e japonês); 15 artigos duplicados; 09 publicações não científicas; 114 artigos com tema diverso da intenção de pesquisa; e 61 artigos cujo estudo se detinha em um único evento de treinamento. Além desses, foram identificados 17 artigos teóricos e de revisão de literatura, analisados separadamente. Por fim, 01 artigo

encontrado em pesquisa livre no periódico *Cadernos de Finanças Públicas*, não indexado em qualquer das bases, foi adicionado por sua relevância perante a intenção de pesquisa. O levantamento encerrou-se com um total final de 61 artigos científicos.

O delineamento metodológico da amostra final resultou em uma proporção de 65% de estudos com abordagem quantitativa, 18% com abordagem qualitativa e uma parcela de 16% em que as duas abordagens são utilizadas. A coleta de dados secundários em bases de dados é o método mais prevalente, utilizado em 25 artigos (40%), seguido dos *surveys*, utilizados em 23 artigos (38%). Com relação aos métodos de análise, de forma geral, considerando-se todas as pesquisas, os mais utilizados na amostra são a Análise Envoltória de Dados (22, 36%), Análise de Conteúdo (14, 23%), Estatística Descritiva (12, 21%) e Estatística Inferencial (11, 20%). Além dessa distinção metodológica, as variáveis operacionalizadas nos estudos também se distinguem bastante, devido às diferentes perspectivas de análise do treinamento. Com base na análise dos objetivos de pesquisa, 22 artigos (37%) enfatizavam variáveis critério de aprendizagem, 34 artigos (62%) variáveis de resultados ou processos organizacionais e 05 (08%) versavam sobre análises no contexto de políticas públicas. Após a seleção dos artigos com base em parte das etapas previstas pela *Methodi Ordinatio* e após a sua leitura, a seção seguinte tratará das variáveis utilizadas nos artigos empíricos selecionados.

2.5.1 Análise das variáveis utilizadas nos artigos empíricos selecionados

Como principal objetivo da presente revisão de literatura, a leitura integral dos 61 artigos visou a identificação das variáveis expressamente abrangidas por cada estudo. Para estudos qualitativos, considerou-se como incidência a descrição ou análise da respectiva variável no capítulo de resultados. A categorização utilizou como base a estruturação do modelo para avaliação da eficiência em TD&E o qual se baseou na estrutura avaliativa do CIPP (Stufflebeam, 2003) e dos modelos lógicos (McLaughlin & Jordan, 2015). Dessa forma, as variáveis serão distribuídas conforme as seguintes categorias: 1) Contexto, 2) Insumo, 3) Processo, 4) Produto, 5) Resultados para o Indivíduo (resultados de curto e médio prazo), e 6) Resultados para a organização e para a sociedade (longo prazo). Assim, a categorização se fundamentou nos pressupostos teóricos de quais variáveis constituem tais componentes estruturantes das avaliações em TD&E. A Tabela 9 apresenta as variáveis encontradas referentes a cada categoria, bem como a quantidade e proporção de artigos que operacionalizaram tais indicadores e medidas.

Tabela 9
Variáveis encontradas na revisão de literatura

<i>Variáveis</i>		<i>Qtd. Artigos</i>	<i>% Artigos</i>
<i>Variáveis de Contexto</i>			
1	Indicadores da força de trabalho	26	43%
2	Estrutura, setor e localização	23	38%
3	Contexto socioeconômico	3	5%
4	Suporte Organizacional	1	2%
<i>Variáveis de Insumo (Instrucional)</i>			
4	Características da clientela: dados sociodemográficos, profissionais, crenças e percepções.	24	40%
5	Características de instrutores e professores	16	26%
<i>Variáveis de Insumo (Recursos)</i>			
6	Características da equipe administrativa do treinamento	8	13%
7	Logística, Instalações e suporte técnico para o Treinamento	9	15%
8	Investimento em Treinamento	31	51%
<i>Variáveis de Processo</i>			
9	Procedimentos e características instrucionais dos Treinamentos	26	43%
10	Implementação de Treinamento	22	36%
<i>Variáveis de Produto</i>			
11	Tempo de Treinamento	16	26%
12	Quantidade de Treinamento	12	20%
13	Participações em Treinamento.	19	32%
<i>Variáveis de Resultado Individual</i>			
14	Reação e satisfação com os treinamentos	14	24%
15	Aprendizagem com treinamentos	14	24%
16	Percepção de impacto de treinamentos	9	15%
17	Desempenho individual no trabalho	7	12%
18	Empregos obtidos	13	22%
19	Salários obtidos	12	20%
<i>Variáveis de Resultado Organizacional</i>			
20	Medidas e resultados financeiros	15	25%
21	Mudanças organizacionais	23	38%
22	Resultados e desempenho organizacional	20	33%
22	Resultados para a força de trabalho	8	13%
23	Resultados para a sociedade	1	2%

Nota. Fonte: Elaboração própria.

As variáveis de contexto referem-se, de forma, geral a características organizacionais e ambientais considerados na análise dos processos de treinamento e educação. Quatro categorias de variáveis foram encontradas: aquelas relativas à força de trabalho, compreendem principalmente aspectos quantitativos como o porte da organização – número de funcionários – proporções de cargos, dados demográficos e profissionais; características da organização e de sua estrutura, envolvendo o tipo de organização (e.g., pública, privada, nacional, multinacional, etc.),

setor de atividade da indústria, localização, número de unidades descentralizadas ou filiais, a existência de conselhos de trabalho, a existência de departamentos de gestão de pessoas, unidades de treinamento ou universidades corporativas, existência ou intensidade do uso de determinadas tecnologias nas atividades organizacionais, tempo de existência e características da alta direção da organização, Por fim, três artigos consideram em suas análises aspectos socioeconômicos como dados demográficos e sobre estruturas familiares, produto interno bruto, dados do mercado de trabalho, custo de vida e estratificação social. Apenas um artigo buscou mensurar suporte organizacional ao treinamento (Odelius & Siqueira Júnior, 2011).

Com relação às variáveis de insumo em perspectiva instrucional, classificação proposta nesse estudo, as principais variáveis se referem às características da clientela de treinamentos e atividades educacionais, assim como às características de instrutores e professores. Tais variáveis são abrangidas, respectivamente em 40% e 26% dos artigos. Os dados da clientela envolvem cargos, salários, perfil sociodemográfico, escolaridade e nível de treinamento prévio e percepções e expectativas com relação a treinamentos. Os dados de instrutores/professores são similares, acrescentando-se a formação, desempenho acadêmico e como instrutor, carga horária, e experiência.

As variáveis de insumos relativas a recursos representam os recursos financeiros, humanos e materiais, sendo a primeira a mais operacionalizada na amostra de artigos, em que 51% deles consideram medidas de investimento no treinamento operacionalizadas de diversas formas. Uma quantidade menor de artigos considera variáveis relativas aos recursos materiais, abrangendo infraestrutura para treinamento e equipamentos (15%), e variáveis relativas à equipe de treinamento, em sua maioria o quantitativo de pessoal dedicado à unidade de treinamento o apoio a programas educacionais.

Com relação às variáveis que versam sobre os processos de treinamento, cabe mencionar que os estudos não adotavam uma perspectiva formativa. Parte das variáveis utilizadas em 43% dos estudos considerava características instrucionais dos cursos e programas, abrangendo seus conteúdos, modalidade, tipo de curso, se eram internos ou externos, carga horária, complexidade, adequação dos objetivos e do planejamento instrucional, métodos de ensino, materiais didáticos e seu alinhamento a necessidades de trabalho. Outra categoria de variáveis abrangidas por 36% dos artigos se voltava para os processos e programas de treinamento, averiguando se havia treinamento na organização, com mais estudos analisando se haviam políticas e avaliações formais de treinamento na organização. Duas variáveis também frequentes eram a análise de alinhamento estratégico (à organização) e legal (às normas de governo) dos programas de treinamento, seguidas da existência de ANT, procedimentos de formação de instrutores, cursos de ambientação ou formação inicial e mapeamento de competências.

Considerando-se a categorias produtos, como aquilo que é imediatamente resultante das atividades de um programa, encontraram-se três categorias de variáveis, a primeira relativa a medidas de tempo de treinamento em 26% dos artigos, com dados e índices de horas ou dias de treinamento. Variáveis sobre número de cursos realizados eram utilizados em 20% dos artigos e variáveis sobre quantitativos e índices de participação ou número de treinados/formados em 32%.

Os resultados mensurados no nível dos indivíduos envolvem medidas nos níveis tradicionais do treinamento de reação (24%), aprendizagem (24%), percepção de desempenho (15%) e desempenho mensurado com medidas objetivas (7%). Duas medidas bastante operacionalizadas na literatura de educação também estavam presentes nessa categoria, relativa aos empregos (13%) e salários (12%) por vezes mensurados, antes e após a ação educacional, e quase sempre de forma conjunta. Com relação aos resultados mensurados no nível da organização parte dos artigos utilizou variáveis de desempenho financeiro (25%), mudanças organizacionais (38%), tais como melhoria da eficiência, produtividade, inovações, medidas específicas de resultados (33%), tais como aumento do número e satisfação de clientes, quantidade de vendas, participação de mercado, reputação e desempenho acadêmico da instituição. Além disso, alguns artigos (13%) consideraram resultados para a força de trabalho, como a melhoria de clima organizacional e satisfação no trabalho e redução de *turnover*. Por fim, um artigo previu efeitos dos programas de treinamento para o contexto social (Meneses & Abbad, 2009).

A estrutura de classificação proposta, portanto, comporta as variáveis encontradas termina por explicitar os componentes e dimensões de processos de treinamento e educacionais mais abrangidas. Pode-se concluir que há uma distribuição equilibrada entre abordagens cujos critérios de resultado estão centrados nos indivíduos e nas organizações. A maior incidência de artigos que consideram medidas de investimento em treinamento reflete a perspectiva da eficiência, utilizada como critério para a inclusão dos artigos na revisão, sendo as variáveis para esse tipo de recurso mais recorrente que as demais, relativos a recursos humanos e materiais. Deve-se ter em consideração que há grande variedade de objetivos e metodologias dos artigos, de forma que nem sempre a análise da eficiência era o principal objetivo do estudo. De forma geral, os estudos que tiveram essa perspectiva como objetivo de análise foram estudos econométricos, dentre eles os que utilizaram a análise envoltória de dados. Para compreensão de como essas variáveis foram operacionalizadas com esse método, as próximas seções se concentrarão nos 22 artigos da amostra que utilizaram essa abordagem.

2.5.2 Artigos com aplicação do DEA

Como um dos objetivos secundários da presente análise, foram encontrados 22 artigos que utilizaram o DEA para avaliação de treinamentos ou, como na maioria dos casos, instituições e programas educacionais, mais relacionados ao campo de educação superior ou profissional. A Tabela 10 elenca as DMUs focalizadas em cada um desses estudos.

Tabela 10

DMUs de artigos com abordagem DEA

<i>Tipo de DMU</i>	<i>Frequência</i>
Faculdades	4
Instituições de Ensino Profissionalizante	4
Escolas de Negócios	3
Instituições de Ensino Superior	2
Programas de <i>Master in Business Administration</i> (MBA)	5
Programas de Treinamento Organizacional	1
Organização Pública	1
Empresas Privadas	1
Indústria	1
<i>Total</i>	22

Nota. Fonte: Elaboração própria.

As primeiras cinco categorias, portanto, referem-se a estudos com ênfase na eficiência educacional. Os demais quatro estudos referem-se à avaliação do treinamento, propriamente dito. O estudo de Salcedo (2020) avalia a eficiência e desempenho de um programa de formação de professores em sete *campi* de uma determinada universidade pública nas Filipinas. Por meio da aplicação de cinco modelos DEA com perspectivas de resultados para professores, alunos, *status* do programa, gestão universitária e pesquisa, o estudo operacionaliza múltiplas variáveis, considerando como *inputs* variáveis do programa de treinamento e sua eficiência na produção de *outputs* tais como, melhor desempenho de professores, maior aprendizagem e número de alunos formados, percepção de desempenho administrativo de unidades acadêmicas, melhoria de avaliações institucionais e produção científica (Salcedo, 2020). O autor então compila os índices de eficiência, visando criar um índice geral de desempenho universitário.

O estudo de Silva et al. (2016) já apresentado representa o único estudo a utilizar o DEA para avaliação de treinamento de servidores públicos administrativos. Não obstante, similar ao estudo de Salcedo (2020), realiza o estudo em *campi* de uma instituição de ensino, o que aproxima ambos os estudos da linha tradicional de aplicação do DEA em contextos educacionais. O estudo de Mehra, Langer, Bapna & Gopal (2014) pode ser considerado um dos poucos estudos que visou avaliar o impacto do treinamento em medidas financeiras da organização, no entanto, alinha-se ao campo da economia, remetendo ao comentado por Aguinis & Kraiger (2009) e Borges-

Andrade (2002) de que a maioria dos estudos de efetividade termina por ser realizado por esse campo. A partir da perspectiva da contribuição do capital humano, os autores utilizam um modelo econométrico para desenvolvimento de uma medida de ROI do treinamento em firmas de tecnologia da informação. Os autores utilizam o DEA como parte do modelo econométrico para geração de um parâmetro de eficiência do treinamento visando estimar seu retorno com base nos gastos globais com treinamento, quantidade de empregados, receita operacional e valor dos ativos da organização.

O artigo de Ng & Li (2003) também foge à tradicional perspectiva educacional na aplicação do DEA, no entanto não avalia a eficiência do treinamento em si, mas o seu efeito na eficiência organizacional da indústria com base em dados primários coletados por meio de um *survey* junto a 515 indústrias manufatureiras na China, podendo ser considerado um estudo de eficiência organizacional com integração do treinamento ao modelo. Os autores coletam múltiplos dados objetivos (*hard data*) acerca de dados da força de trabalho, dos resultados financeiros e operacionais da organização, e da provisão de treinamento. Foram levantados a quantidade de dias e o investimento dedicados a treinamento (intensidade) – discriminando-o entre técnico e gerencial – e o percentual de pessoal, discriminados por funções e níveis hierárquicos, além de medidas de controle relativas às características da organização. Os autores utilizam as quantidades de operários, técnicos e gerentes treinados, juntamente com uma medida dos ativos das empresas como *input* para a geração de vendas, sendo este o *output* único. Complementando o modelo DEA com o modelo TOBIT de regressão, os autores encontram efeitos positivos do treinamento para os níveis técnico e gerencial, sendo o maior efeito devido a treinamentos para pessoal de cargos técnicos na organização.

Por fim, apesar de focar DMUs com propósito educacional, o artigo de Cubbin & Zamani (1996) relaciona-se mais diretamente ao campo de avaliação de programas públicos, sendo o primeiro da série histórica da amostra. Assim, dado o seu relativo pioneirismo, os autores adotam um ponto de vista metodológico com o objetivo de analisar a adequação do DEA para avaliação do desempenho de uma política pública britânica voltada para o fornecimento de treinamentos em organizações. A pesquisa avalia o desempenho de instituições conhecidas como *Training and Enterprise Councils* - TECs⁶, entes privados concessionários da política governamental, que contratam e ofertam treinamentos para jovens trabalhadores nas empresas. Baseando-se em dados de 75 TECs, resgatados em bases documentais do Departamento do

⁶ Conselhos de Treinamento e Empresa – tradução livre

Trabalho (*Department of Employment*), Cubbins e Zamani (1996) propõem modelos econométricos para análise da eficiência desses conselhos, definindo insumos (*inputs*), relacionados aos custos dos treinamentos, e produtos (*outputs*) representativos da qualidade do conteúdo e quantitativo de trabalhadores treinados. Também são propostas variáveis contextuais socioeconômicas – em DEA considerados como fatores não-discricionários (Ferreira & Gomes, 2020) – que devem ser levados em consideração na análise para fins de controle. Os autores recomendam a aplicação do DEA em função de sua robustez, após contrastar os resultados de eficiência obtidos com esse modelo com o de outras técnicas de análise, como a análise de regressão, modelos de fronteira estocástica (SFA) e fronteira determinística.

Os estudos acima citados foram aqueles que mais se aproximaram de um conceito clássico de treinamento enquanto atividade voltada para a melhoria da atividade organizacional. Uma parte significativa dos demais estudos, no entanto, tem como perspectiva de análise a eficiência e desempenho da educação superior e profissional em salários e empregabilidade no nível do indivíduo. Visando proporcionar uma visão geral dos 22 estudos que utilizaram o DEA, a Tabela 11 apresenta os produtos considerados e as respectivas referências, de forma a se traçar os tipos de resultados normalmente enfatizados.

Tabela 11

Principais produtos (outputs) utilizados nos modelos DEA da amostra

Produtos	Frequência	Estudos
Empregabilidade/Empregos adquiridos por egressos	10	Haksever e Muragishi (1998); Aviles-Sacoto, Cook, Imanirad e Zhu (2015); Fisher, Kiang e Fisher (2007); Hirao (2012); Hsu, James e Chao (2009); Salcedo (2020); Fisher, Chi, Fisher e Kiang (2017); Pai, Khan e Mukherjee (2019); Bogetoft e Wittrup (2017); Kao, Chiu e Ko (2014)
Salários Adquiridos por egressos	8	Haksever e Muragishi (1998); Debnath e Shankar (2009); Fisher, Kiang e Fisher (2007); Hirao (2012); Hsu, James e Chao (2009); Ekiz e Tuncer (2020); Kumar (2019); Fisher, Chi, Fisher e Kiang (2017)
Quantidade de pessoal treinado/formado	8	Kumar (2019); Nkonki, Ntlabathi e Ncanywa (2014); Salcedo (2020); Bogetoft e Wittrup (2017); Kao, Chiu e Ko (2014); Silva, Diniz e Roratto (2016); Talukder (2011); Cubin e Zamani (1996)

Reação e satisfação com curso	5	Debnath e Shankar (2009); Hsu, James e Chao (2009); Ekiz e Tuncer (2020); Nojavan, Heidari e Mohammaditabar (2020); Debnath, Shankar e Kumar (2008)
Aprendizagem	4	Kong e Fu (2012); Salcedo (2020); Pai, Khan e Mukherjee (2019); Bogetoft e Wittrup (2017)
Desempenho Organizacional (e.g. vendas, receita)	3	Mehra, Langer, Bapna e Gopal (2014); Kumar (2019); Ng e Li (2003)
Reputação Organizacional	3	Haksever e Muragishi (1998); Kumar (2019); Fisher, Chi, Fisher e Kiang (2017)
Desempenho acadêmico da instituição	3	Kumar (2019); Salcedo (2020); Fisher, Chi, Fisher e Kiang (2017)
Posição no Mercado	2	Salcedo (2020); Fisher, Chi, Fisher e Kiang (2017)
Desempenho individual	2	Salcedo (2020); Silva, Diniz e Roratto (2016)
Percepções da alta gerência	2	Debnath e Shankar (2009); Debnath, Shankar e Kumar (2008)
Outros: <i>turnover</i> , equidade, comprometimento, atividades extracurriculares, fatores ambientais, contribuição para o produto interno bruto	6	Bogetoft e Wittrup (2017); Ekiz e Tuncer (2020); Debnath, Shankar e Kumar (2008); Salcedo (2020); Talukder (2011); Cubin e Zamani (1996)

Nota. Fonte: Elaboração própria.

Observa-se com base na Tabela 11 a prevalência, portanto, de estudos com ênfase em análise da eficiência e desempenho de ações e programas educacionais em empregos (10) e salários (8), representando uma abordagem comum para a avaliação de instituições de ensino (Brauchle & Schmidt, 2004). Dentre as variáveis-critério comuns aos campos de TD&E e Educação há a consideração da quantidade de pessoal treinado ou formado (8) e aspectos relacionados aos tradicionais resultados no nível individual: reação (5), aprendizagem (4) e desempenho (2).

O nível organizacional é abrangido por variáveis de produto que denotam a eficiência organizacional com a consideração do TD&E como variável para a melhoria do desempenho geral da organização, em termos de vendas e receitas (3); *status* e reputação, baseados em classificações e opiniões de *stakeholders* (3); posição no mercado (2); mudanças nas percepções da alta gerência (2); e desempenho acadêmico, para o caso de instituições de ensino (3). Por fim, algumas medidas como *turnover*, equidade entre homens e mulheres no trabalho, comprometimento com a profissão, aumento de atividades extracurriculares de estudantes, e contribuição para o produto interno bruto – como efeito da educação – são alguns dos produtos

e efeitos menos operacionalizados. Com relação aos insumos, a Tabela 12 apresenta quais fatores são operacionalizados como capazes de gerar os produtos acima discriminados.

Tabela 12

Principais insumos utilizados nos modelos DEA da amostra

Insumos	Frequência	Estudos
Quantidade total de funcionários	9	Mehra, Langer, Bapna e Gopal (2014); Kumar (2019); Nojavan, Heidari e Mohammaditabar (2020); Salcedo (2020); Pai, Khan e Mukherjee (2019); Bogetoft e Wittrup (2017); Kao, Chiu e Ko (2014); Ng e Li (2003); Talukder (2011)
Orçamento ou receita para treinamento/educação	7	Haksever e Muragishi (1998); Debnath e Shankar (2009); Fisher, Kiang e Fisher (2007); Hsu, James e Chao (2009); Fisher, Chi, Fisher e Kiang (2017); Talukder (2011) Debnath, Shankar e Kumar (2008)
Desempenho Escolar Progresso	7	Haksever e Muragishi (1998); Aviles-Sacoto, Cook, Imanirad e Zhu (2015); Fisher, Kiang e Fisher (2007); Hirao (2012); Hsu, James e Chao (2009); Salcedo (2020); Fisher, Chi, Fisher e Kiang (2017);
Quantidade de instrutores e professores	7	Haksever e Muragishi (1998); Debnath e Shankar (2009); Kumar (2019); Nojavan, Heidari e Mohammaditabar (2020); Nkonki, Ntlabathi e Ncanywa (2014); Salcedo (2020); Kao, Chiu e Ko (2014)
Quantidade de participantes, estudantes, matrículas	5	Aviles-Sacoto, Cook, Imanirad e Zhu (2015); Debnath e Shankar (2009); Nkonki, Ntlabathi e Ncanywa (2014); Salcedo (2020); Bogetoft e Wittrup (2017)
Horas/Tempo de treinamento	4	Fisher, Kiang e Fisher (2007); Salcedo (2020); Pai, Khan e Mukherjee (2019); Kao, Chiu e Ko (2014)
Infraestrutura para treinamento/educação	4	Nojavan, Heidari e Mohammaditabar (2020); Bogetoft e Wittrup (2017); Ng e Li (2003); Debnath, Shankar e Kumar (2008)
Equipamentos para treinamento/educação	4	Kumar (2019); Bogetoft e Wittrup (2017); Ng e Li (2003); Talukder (2011)
Custos totais	4	Kumar (2019); Nojavan, Heidari e Mohammaditabar (2020); Bogetoft e Wittrup (2017); Cubin e Zamani (1996)

Tempo de Experiência de funcionários	3	Haksever e Muragishi (1998); Aviles-Sacoto, Cook, Imanirad e Zhu (2015); Salcedo (2020)
Índice de aceitação de candidaturas (Universidades)	2	Haksever e Muragishi (1998); Aviles-Sacoto, Cook, Imanirad e Zhu (2015)
Quantidade de programas de treinamento/educação	2	Salcedo (2020); Debnath, Shankar e Kumar (2008)
Formação e qualificação de instrutores (Treinamento)	2	Haksever e Muragishi (1998); Salcedo (2020)
Escolaridade dos participantes/estudantes	2	Salcedo (2020); Kao, Chiu e Ko (2014)
Custos Operacionais	2	Kumar (2019); Bogetoft e Wittrup (2017)
Custo dos treinamentos	2	Mehra, Langer, Bapna e Gopal (2014); Silva, Diniz e Roratto (2016)
Custo de funcionários	2	Nkonki, Ntlabathi e Ncanywa (2014); Bogetoft e Wittrup (2017)
Insumo singular	2	Ekiz & Sakar, 2020; Kong & Fu, 2012
Outros: Formação, qualificação e perfil dos funcionários, conteúdo e complexidade do treinamento, custo de viagens, hospedagens e materiais, reputação, custo por estudante, idade e salário anterior dos estudantes, quantidade de ações educacionais, idade da DMU, quantidade de sindicatos reconhecidos	8	Salcedo (2020); Nkonki, Ntlabathi e Ncanywa (2014); Kumar (2019); Fisher, Kiang e Fisher (2007); Hirao (2012); Bogetoft e Wittrup (2017); Haksever e Muragishi (1998); Hsu, James e Chao (2009);

Nota. Fonte: Elaboração própria.

A análise da Tabela 12 demonstra que boa parte dos *inputs* se referem a medidas de produtividade e de recursos, como recursos humanos, materiais, infraestrutura e diferentes dimensões de custo. Os recursos humanos se traduzem em termos da quantidade total de funcionários e educadores (instrutores e professores) à disposição da organização, sendo que também podem ser consideradas características dos indivíduos, tais como níveis de qualificação. Quando se almeja utilizar o DEA apenas para fins de comparação de desempenho, pode-se prescindir da definição de insumos adotando-se uma medida univariada de insumo singular (*Single Input*) como foi realizado em dois dos estudos da amostra (Ekiz & Sakar, 2020; Kong & Fu, 2012).

A despeito da grande quantidade de variáveis, os modelos observados na amostra adotam em média entre 3 e 4 variáveis de *input* e de *output*, em um total de no máximo, 6 a 7 variáveis, em média, similar ao observado na revisão realizada por Suescun et al. (2007) no campo da educação superior. Dentre os modelos, dez deles expressamente adotam a orientação para os *outputs*, isto é, buscam avaliar a eficiência em termos de margens para aumento de desempenho, refletindo, possivelmente, o fato de que os insumos não sejam facilmente gerenciáveis ou ajustáveis. Não obstante, seis estudos definem sua orientação para o *input*,

visando à economia de recursos, três estudos definem uma orientação tanto para *inputs* quanto *outputs* e três deles não fornecem essa informação em seus capítulos de método. A esse respeito, oito dos estudos utilizam o modelo básico CCR de Charnes et al. (1978), implicando a premissa de retornos constantes nas atividades analisadas, cinco utilizam o modelo BCC de Banker et al. (1984) para retornos variáveis, quatro combinam e comparam as duas perspectivas e um descreve um modelo específico mais recente do DEA (Order-M) (i.e. Silva et al., 2016). Cinco estudos não informam qual o modelo utilizado. Não obstante, alguns artigos utilizam versões mais avançadas ou adaptadas do DEA, tais como as análises de dados em painel (longitudinal) com cálculo de índice Malmquist (2), que mensura a evolução temporal da eficiência e da fronteira de produção da amostra; regressões com fins à análise de relações de causalidade (Ferreira & Gomes, 2020); modelos *fuzzy* (1) para sistemas não-lineares, modelos DEA de múltiplos estágios, tais como o *network* DEA (1) e, por fim, modelos econométricos aperfeiçoados, adaptados às condições específicas de análise (3).

A presente seção visou investigar a literatura empírica para definição de fundamentos e referências para o desenvolvimento de um modelo de avaliação da eficiência de programas de treinamento. Em sua maioria, os artigos encontrados não analisam a eficiência do gasto em TD&E visando apontar alternativas de alocação dos investimentos para melhoria da provisão de treinamento. A maior parte dos estudos encontrados se voltam para a análise da eficiência educacional, ou utilizam o treinamento como variável que contribui para a eficiência de atividades organizacionais. Além disso, apesar da operacionalização de variáveis de reação e satisfação com o curso, aprendizagem e melhoria de desempenho, os estudos que utilizam o DEA não se fundamentam na literatura de TD&E. Apesar dessa aparente proximidade entre o TD&E e as áreas em que o DEA vem sendo aplicado, as perspectivas que orientam esses estudos normalmente se fundamentam nos campos da educação e da economia. O estudo de Silva et al. (2016), consistiu no único estudo encontrado que utiliza um modelo de avaliação de treinamento como referência para a aplicação do DEA, no caso o modelo IMPACT (Abbad, 1999), em que utiliza como variável de produto medidas de reação de servidores em uma organização pública. Com isso, demonstra-se que apesar de rara, há a possibilidade de articulação entre esses dois campos. Além disso, assim como no campo de TD&E, verificou-se que os estudos de eficiência adotam perspectivas e variáveis-critério similares ao TD&E, podendo focar resultados aos indivíduos, às organizações, ao ambiente ou aspectos do próprio processo de treinamento.

Conclusivamente ao presente referencial teórico, verificou-se, primeiramente, que os modelos de avaliação de TD&E possuem como perspectiva principal a eficácia e efetividade

de treinamentos. Conforme Suchman (1967) a perspectiva de avaliação da eficiência pode apresentar contribuições para a melhor gestão dos recursos do treinamento. Como visto em estudos da literatura de avaliação educacional, podem ser suscitadas melhorias em programas e a comparação entre diferentes unidades de treinamento. A literatura empírica de avaliação da eficiência em treinamento no contexto organizacional, apesar de escassa, apresenta variados estudos que consideram como insumos os investimentos em treinamento. No entanto, relaciona-os diretamente aos resultados para os indivíduos, sem consideração das características dos treinamentos. Desta forma, tais estudos não possibilitam a análise para a melhoria de programas de treinamento. Os estudos e variáveis levantadas, no entanto, servirão de apoio para a operacionalização de um modelo de avaliação da eficiência de TD&E com utilização da Análise Envoltória de Dados, visando apontar melhorias ao próprio programa de treinamento.

Para prosseguimento do presente estudo, serão executadas as etapas propostas por Senra et al. (2007); Golany & Roll (1989) e Ferreira e Gomes (2020) para especificação de um modelo de Análise Envoltória de Dados, quais sejam: 1) Seleção por Julgamento, com a realização de pesquisa de campo para conhecimento das atividades realizadas e coleta de dados; 2) Análises estatísticas sobre os dados coletados; e 3) Análises com modelos de DEA para definição de um modelo adequado e parcimonioso para estimativa da eficiência (Rosano-Peña, 2008). Tais etapas visam concretizar o primeiro objetivo específico da presente pesquisa, relativo à definição das variáveis para avaliação da eficiência em TD&E. Haja vista que o cumprimento dessas etapas e do citado objetivo específico ensejam a realização de pesquisa empírica, a sua realização será descrita no capítulo seguinte referente ao método.

3 Método

A presente pesquisa tem como objetivo a proposição de um método para a avaliação da eficiência de unidades de treinamento. A perspectiva de análise tem como foco as decisões para gestão dos recursos de treinamento, visando a determinação da eficiência de sua utilização para a provisão de ações educacionais em contextos organizacionais. Para tanto, buscou-se a definição de um modelo de Análise Envoltória de Dados e a consequente identificação de variáveis relevantes que permitam a análise comparativa de unidades de treinamento. Diante disso, esta seção, primeiramente, descreve as características do contexto de pesquisa selecionado e o delineamento metodológico do estudo proposto. Em seguida, apresenta-se a pesquisa de campo preliminar realizada com o intuito de compreender como são realizados os processos de treinamento e quais variáveis podem ser utilizadas para determinação da sua eficiência. A partir da definição dessas variáveis, foi possível a coleta dos dados necessários para a avaliação, de forma que são então descritos os instrumentos desenvolvidos, assim como os procedimentos de coleta de dados e as organizações componentes da amostra final. Por fim, são abordadas as etapas relativas à análise dos dados para proposição do modelo DEA definitivo adotado.

3.1 Universo de Pesquisa

Para realização do processo de seleção das organizações que compuseram a amostra de DMUs foram considerados os requisitos exigidos pela Análise Envoltória de Dados. Primeiramente, as organizações selecionadas e suas respectivas unidades de treinamento devem ser homogêneas para que sejam comparáveis, no sentido de que devem atuar em condições e contextos similares e realizar as mesmas atividades com os mesmos tipos de recursos (Rosano-Peña, 2008; Mariano, 2007). Nesse sentido, as organizações públicas apresentam-se como um campo adequado para a realização da presente pesquisa uma vez que, conforme Bach (2010), tendem a ser regidas por leis definidas por um órgão central e adotar práticas similares estabelecidas em normas comuns, seja em relação à gestão orçamentária, procedimentos de aquisição e implementação de ações educacionais, assim como apresentam certa homogeneidade em relação às políticas de gestão de pessoas de forma geral. Dessa forma, adotou-se como universo de pesquisa as organizações integrantes da Administração Pública Federal.

Todas as organizações integrantes da Administração Pública Federal possuem em sua estrutura unidades incumbidas de executar as atividades de gestão, capacitação e

desenvolvimento de servidores. Tais unidades vinculam-se ao Sistema de Pessoal Civil da Administração Federal – SIPEC e são gerenciadas e monitoradas pelo órgão central do SIPEC, atualmente, a Secretaria de Gestão e Desempenho de Pessoal do Ministério da Economia – SGP/ME. Dessa forma, o universo amostral do presente estudo constitui-se das unidades integrantes do SIPEC nos órgãos da Administração Pública Federal Direta, Indireta, Autárquica e Fundacional. Nesse contexto, as atividades de Treinamento são regidas pela Política Nacional de Desenvolvimento de Pessoas (PNDP) que define como principais objetivos para essas unidades o desenvolvimento permanente de servidores público de forma alinhada às estratégias organizacionais e de governo, assim como a provisão de ações educacionais de maneira eficiente e de forma equânime. Conforme dados do Painel Estatístico de Pessoal⁷ do Governo Federal, tais unidades totalizaram 191 órgãos em dezembro de 2019, sendo responsáveis pela gestão e qualificação de cerca de 1.273.584 servidores públicos federais estatutários.

3.2 Caracterização da pesquisa

O estudo proposto se caracteriza como uma pesquisa teórico-empírica, uma vez que se baseia na coleta e análise de dados e informações acerca do campo de estudo visando identificar, compreender e descrever a sua realidade. A pesquisa possui natureza descritiva, com relação aos seus fins, uma vez que aborda o fenômeno por meio da identificação e mensuração de seus componentes visando a descrição de suas características (Collis & Hussey, 2005). Para tanto, adotou uma abordagem quantitativa visando estabelecer a compreensão do fenômeno a partir de medidas e dados numéricos descritivos de suas variáveis constituintes. De forma complementar, a pesquisa adotou uma abordagem qualitativa, primeiramente para a realização de pesquisa preliminar de campo visando à compreensão do fenômeno estudado e, posteriormente, para a discussão dos índices de eficiência das organizações estudadas com as características dos processos de treinamento realizados nas organizações.

A definição de uma abordagem predominantemente quantitativa se justifica em razão do objetivo de pesquisa, qual seja a proposição de um método de avaliação da eficiência por meio da Análise Envoltória de Dados (*Data Envelopment Analysis – DEA*), uma técnica matemática de programação linear que permite o cálculo de índices de eficiência de unidades produtivas por meio da comparação de seus processos de geração de determinados produtos a partir de determinados insumos (Barbosa & Fuchigami, 2008; Charnes et al., 1978; Ferreira &

⁷ <https://www.painel.pep.planejamento.gov.br/>

Gomes, 2020; Rosano-Peña, 2008). Conforme Senra et al. (2007) e Ferriera e Gomes (2020), a definição de um modelo DEA deve partir de um conhecimento aprofundado do processo a ser analisado, com base, principalmente, na consulta a especialistas e responsáveis pela sua gestão. Dessa maneira, anteriormente à realização dos procedimentos de pesquisa propostos, realizou-se uma pesquisa preliminar de campo.

3.3 Pesquisa preliminar de campo

Para compreensão das atividades de treinamento executadas pelas organizações públicas federais foi realizada uma breve pesquisa de campo. Para tanto, foram contatados gestores e servidores integrantes de unidades de treinamento de 12 (doze) organizações públicas federais no mês de março do ano de 2020. A pesquisa, de natureza exploratória com abordagem qualitativa, caracterizou-se pela realização de entrevistas não-estruturadas, voltadas para a aplicação de um conjunto de questões acerca de quais eram as principais atividades de treinamento realizadas pelas organizações. O critério para encerramento da amostragem por acessibilidade das organizações foi a saturação dos dados, em que a partir de determinado ponto não se obtém novas informações sobre o objeto investigado (Flick, 2009). Dessa forma, após o contato com as organizações chegou-se a uma definição convergente acerca de como são realizados os processos de TD&E nesse contexto.

Os questionamentos tiveram como fundamento os procedimentos e práticas prescritas para processos de treinamento (Salas et al., 2012; Borges-Andrade & Abbad, 1996; Meneses et al., 2010), tendo sido perguntado às organizações a respeito de suas rotinas para realização de análises de necessidades de treinamento, planejamento, execução e avaliação de treinamento, assim como a consideração de possíveis fatores interferentes previstos em estudos específicos sobre a implementação de políticas de gestão e desenvolvimento de pessoas no setor público federal no Brasil (e.g. Camões & Meneses, 2016; Côrtes & Meneses, 2019). As principais atividades comumente apontadas pelos entrevistados e as denominações e descrições utilizadas foram:

1. **Levantamento ou Avaliação de Necessidades de Treinamento:** consistindo, em geral, apenas no envio de formulários às unidades organizacionais.
2. **Mapeamento de Competências:** consistindo na definição de categorias relativas a conhecimentos, habilidades e atitudes, em geral, sem a análises de lacunas.
3. **Elaboração de Planos de Capacitação:** consistindo, em geral, na explicitação de políticas de capacitação e definição dos temas ou cursos a serem realizados no período de um ano podendo haver a previsão orçamentária para os mesmos.

4. **Execução de cursos**, envolvendo:
 - 4.1. **Cursos Abertos:** contratação de vagas em treinamentos com empresas externas;
 - 4.2. **Cursos de Turma Fechada:** contratação de cursos exclusivos para grupos de servidores da organização;
 - 4.3. **Cursos Internos:** realizados com instrutores da própria organização sem contratação externa;
 - 4.4. **Programas permanentes de treinamento:** programas cíclicos para formação de servidores em atividades específicas (e.g. fiscais);
 - 4.5. **Pós-graduação aberta:** matrícula de servidores em cursos de especialização, mestrado ou doutorado em instituições educacionais externas;
 - 4.6. **Pós-graduação fechada:** contratação de cursos exclusivos de especialização, mestrado e doutorado junto a instituições educacionais externas para uma turma de servidores da organização;
 - 4.7. **Capacitação gerencial:** ações educacionais específicas para a capacitação de gestores, distinguindo-se por envolver métodos para desenvolvimento de competências comportamentais de liderança (e.g. *coaching*);
5. **Licença-capacitação e Afastamento:** autorização de afastamento por período determinado para realização de ações educacionais.
6. **Incentivo Educacional:** custeio por meio de bolsas de ações de desenvolvimento e educação – não necessariamente relacionadas ao trabalho – normalmente relacionadas ao aprendizado de idiomas ou obtenção de nível superior.
7. **Gestão administrativa:** atividades de gestão e controle para a execução orçamentária e financeira, instrução de processos de contratação direta e licitações para contratação de ações educacionais; elaboração de relatórios internos e externos, elaboração de normas de capacitação.
8. **Avaliação de treinamentos:** avaliação de reação para a maioria dos casos e, para alguns treinamentos, avaliações de aprendizagem e impacto.

Com relação a fatores contextuais interferentes, foram citados: 1) a necessidade de sistemas informatizados de capacitação; 2) o apoio da alta gestão por meio da consideração da capacitação como atividade estratégica, com previsão de uma estratégia formal com objetivos estratégicos para a capacitação; e 3) o nível hierárquico da unidade de capacitação (estar próximo à alta gestão); 4) o tamanho da equipe de capacitação; e 5) para alguns casos, as dificuldades de capacitação de servidores alocados em unidades descentralizadas nos Estados.

Com relação aos dados disponíveis sobre as ações educacionais, de forma geral, as organizações possuíam planilhas eletrônicas de controle com as seguintes informações:

- i) **Nome e quantidade de treinamentos;**
- ii) **Organização fornecedora;**
- iii) **Forma de implementação/contratação:** contratação direta, licitação, cursos realizados sem necessidade de contratação (parcerias) e cursos internos.
- iv) **Modalidade:** cursos presenciais, cursos a distância e semipresenciais;

- v) **Tipo de treinamento:** em geral, palestra, evento, curso, educação e incentivo;
- vi) **Área temática:** alguns casos, com categorias livres definidas pela própria unidade em seu plano de capacitação;
- vii) **Carga horária do Treinamento;**
- viii) **Quantidade de participantes do Treinamento;**
- ix) **Custo do Treinamento:** valor da contratação, ou do custeio para cursos internos, com discriminação de despesas com deslocamento, estadia e alimentação em alguns casos); e
- x) **Período de Realização.**

A partir dessa consulta preliminar, as informações levantadas foram utilizadas para definição das variáveis que seriam coletadas junto às organizações, visando ao delineamento de um modelo de avaliação da eficiência.

Tendo em vista o número previsto de variáveis a serem coletadas, suscitou-se as possíveis limitações para aplicação do DEA a um número potencialmente reduzido de organizações (Rosano-Peña, 2008). Visando mitigar esses riscos, bem como possibilitar a comparação entre períodos distintos, foi realizado o levantamento de dados dos programas de treinamento relativos a 05 anos, de forma que cada ano correspondeu a uma DMU específica, o que propiciou a aplicabilidade do DEA. Dessa forma, foi executada a coleta de dados com abrangência dos anos de 2015, 2016, 2017, 2018 e 2019, o que caracteriza o estudo como uma perspectiva longitudinal. Tal opção metodológica visou ainda superar uma das dificuldades para realização de análises no setor público (Rosano-Peña, 2008) correspondente à baixa disponibilidade dos dados e suas diferentes formas de registro (e.g. unidades de medida) pelas unidades analisadas (Ferreira & Gomes, 2020). Conforme observa-se na literatura, prevalecem, em razão dessa limitação, estudos que utilizam bases de governo para aplicação do DEA. No entanto, não havia, no período de realização da presente pesquisa, bases de dados estruturadas para as atividades de TD&E. Com base nas informações levantadas, procedeu-se à confecção de instrumentos de pesquisa. A seção seguinte visa descrever os instrumentos utilizados.

3.4 Instrumentos de pesquisa

Para a presente pesquisa foram elaborados três documentos para o levantamento de dados e informações do campo. Para solicitação dos dados e documentos foi elaborada uma carta-convite na qual a pesquisa foi apresentada e descrita com relação aos seus objetivos, resultados esperados e quais dados seriam solicitados à organização (Apêndice B). Foram solicitadas as planilhas de registro de atividades de treinamento com as variáveis conhecidas na pesquisa preliminar de campo para cada um dos anos de 2015, 2016, 2017, 2018 e 2019. Além

disso, foram solicitados e, em alguns casos coletados em portais oficiais de governo, os seguintes documentos institucionais relativos ao período de 2015 a 2019: Planos estratégicos vigentes no período; Políticas de Treinamento; Planos de Capacitação; e Relatórios de Gestão organizacionais, visando subsidiar a melhor compreensão dos processos de treinamento realizados pelas organizações.

Juntamente com a carta-convite para a pesquisa documental, lançou-se mão da coleta de informações por meio de questionários. Os questionários, enquanto técnica de coleta são utilizados quando se tem o objetivo de produzir dados primários para caracterização e descrição de comportamentos e opiniões dos participantes de uma pesquisa (Collis & Hussey, 2005). Os questionários constituíram-se de dois formulários com questões em sua maioria fechadas com informações sobre a pesquisa e instruções de preenchimento. O primeiro formulário denominado “Formulário de Descrição de Equipe” solicitou os quantitativos de pessoal alocados à unidade de treinamento de 2015 a 2019, estratificando-os com relação ao tipo de vínculo dos colaboradores, assim como o seu nível educacional (Apêndice C).

O segundo questionário denominado “Formulário de Descrição de Processos de Capacitação” (Apêndice D) solicitou a manifestação da organização com relação à realização, em cada ano no período de 2015 a 2019, das seguintes atividades: 1) Avaliação de Necessidades de Treinamento; 2) Mapeamento de Competências; 3) Plano de Capacitação; 4) Avaliação de Reação; 5) Avaliação de Aprendizagem; e 6) Avaliação de Impacto. Julgou-se que aspectos concernentes à execução dos cursos poderiam ser apreendidos por meio da pesquisa documental. Também se questionou se a organização possuía em cada um dos períodos um Plano Estratégico formal, com a definição de objetivos de capacitação e, por fim, se a unidade de treinamento utilizava um sistema informatizado para a gestão da capacitação. Para cada questão, criou-se um campo aberto para manifestação dos respondentes caso desejassem tecer observações sobre a respectiva atividade de treinamento.

3.5 Procedimentos e técnicas de coleta de dados

A carta-convite e os formulários foram enviados por e-mail juntamente à uma apresentação da pesquisa para representantes de 45 organizações públicas federais ao longo de dois meses, novembro de 2020 e dezembro de 2020. Os documentos foram enviados para os e-mails institucionais de unidades de gestão de pessoas ou de treinamento, ou diretamente para os respectivos gestores dessas unidades. De forma distinta da estratégia utilizada na etapa preliminar de coleta, o *home office* adotado nas organizações públicas federais durante o período de coleta de dados ensejou maior tempo para os procedimentos de coleta e baixas taxas

de resposta. Sendo assim, do total de 45 organizações contatadas, 16 delas concordaram em participar da pesquisa, tendo enviado os respectivos documentos solicitados. Tais documentos apresentaram diferentes formas de registro e graus integridade das informações solicitadas e, após o período de coleta supramencionado, constatou-se que 10 organizações possuíam os dados necessários para a análise. Os documentos dessas organizações foram submetidos a um processo de tratamento e análise para padronização dos dados. Nesse ínterim, foram solicitadas a todas as organizações complementações, correções e esclarecimentos sobre as informações prestadas, estendendo o processo de coleta até o mês de março de 2021. Para alguns casos de dados faltantes, optou-se pela coleta da respectiva informação em portais oficiais de governo, como o Portal da Transparência, sobretudo para obtenção de dados sobre despesas com ações educacionais. Ao final do processo de tratamento foi possível a consolidação de bases de dados de 06 organizações com as seguintes informações:

- Curso realizado;
- Organização fornecedora do curso;
- Carga Horária;
- Código de Matrícula dos Participantes;
- Período de Realização;
- Local de Realização;
- Modalidade;
- Custo do Treinamento;
- Custos com diárias e passagens;
- Tipo de Ação Educacional.

Além dos dados supracitados, as organizações apresentaram, de maneiras distintas, outros dados, relativos às classificações temáticas dos cursos, *status* de conclusão, previsão no plano de capacitação, formas de contratação, entre outras informações voltadas para a gestão e controle das suas atividades.

3.6 Amostra de organizações pesquisadas

A amostra final de 06 organizações buscou reunir unidades cuja natureza das atividades, assim como a integridade e o padrão de registro dos dados permitisse a comparação. Tendo em vista tais critérios, o processo de amostragem se caracterizou como não-probabilístico por conveniência, uma vez que se baseou principalmente na disponibilidade das organizações para participação na pesquisa e no fornecimento dos dados requeridos pela análise (Shaughnessy, Zechmeister & Zechmeister, 2012). Por esse motivo, o escopo e amostra final de pesquisa

decorreram dos resultados do processo de coleta e tratamento de dados, tendo se constituído de seis organizações públicas federais, sendo 05 Autarquias Especiais e 01 Ministério (Tabela 13).

Tabela 13

Organizações selecionadas

Organização	Classificação
Organização A	Autarquia Especial
Organização B	Autarquia Especial
Organização C	Autarquia Especial
Organização D	Autarquia Especial
Organização E	Autarquia Especial
Organização F	Ministério

Nota. Fonte: elaboração própria.

As organizações selecionadas foram identificadas como Organizações “A”, “B”, “C”, “D”, “E” e “F”. Tendo em vista que foram obtidos dados de treinamentos realizados nos anos de 2015, 2016, 2017, 2018 e 2019, constituíram-se cinco períodos de análise, de forma que, com a utilização do modelo de análise intertemporal, cada organização constituiu-se como uma DMU, ou um caso independente em cada período (Ferreira & Gomes, 2020) perfazendo um total de 30 casos, os quais serão identificados com o final “5” para o ano de 2015, “6” para 2016 e assim por diante, de forma que as organizações serão descritas nas tabelas de resultados como A5, A6... A9, B5, B6... B9, etc. A partir da consolidação de uma base comum de dados para as 06 organizações com abrangência do período de 2015 a 2019, procedeu-se à análise para a seleção definitiva das variáveis constituintes do modelo de Análise Envoltória de Dados

3.7 Definição do modelo e variáveis de análise.

Um dos requisitos para a aplicação do DEA, sob risco de uma estimativa imprecisa dos índices de eficiência, é de que se abranja um número de casos equivalente a, pelo menos, cinco vezes o quantitativo de variáveis analisadas ou um valor equivalente ao produto entre quantidade de variáveis de *input* e de *output* (Ferreira & Gomes, 2020; Rosano-Peña, 2008). Seguindo essa recomendação metodológica, a definição das variáveis que constituiriam o modelo de análise da eficiência levou em consideração a disponibilidade de dados de 30 casos, estabelecendo um limite de 06 variáveis para análise da eficiência.

Considerando dados da revisão de literatura, tal quantitativo enquadra-se no padrão encontrado em outros estudos, que adotam, no máximo, 6 a 7 variáveis para análises no campo da educação e do treinamento (Suescún et al., 2017). Sendo assim, atendo-se ao conceito de eficiência como a relação ótima entre insumos e produtos (Lovell, 1993), buscou-se estreitar esse universo pela seleção de variáveis por julgamento (Senra et al., 2007) e disponibilidade

(Ferreira & Gomes, 2020). Sendo assim, foram consideradas aquelas que melhor possibilitassem o alcance dos objetivos de determinação da eficiência dos gastos de capacitação e abrangência do maior número de servidores, como um critério de equidade, comumente utilizado em pesquisas dessa natureza (Johnes et al., 2017). Por esse motivo, visando abranger essas duas propriedades, definiu-se duas etapas interligadas para representação do processo de treinamento nas organizações públicas federais: 1) Produção e 2) Entrega (Figura 25).

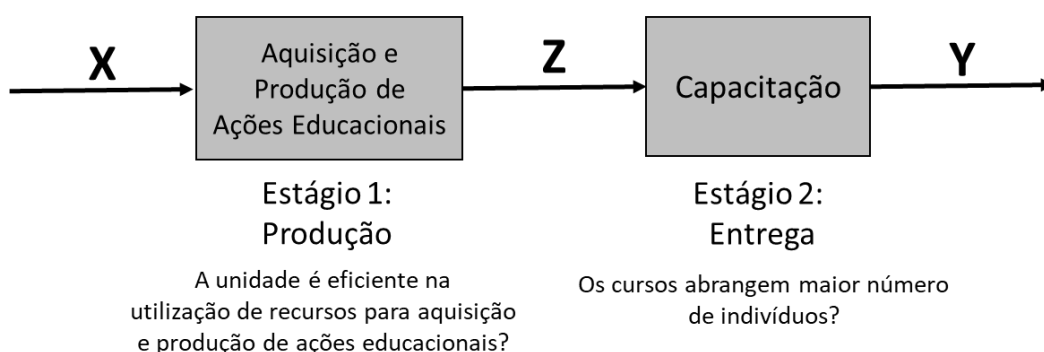


Figura 25. Modelo Conceitual de Análise em dois estágios

Nota. Fonte: Elaboração própria.

Sob essa perspectiva, o modelo define-se como sendo um DEA em rede ou *Network DEA* (Kao, 2017), com duas etapas ou estágios. A interação entre os processos sequenciais é representada por variáveis intermediárias (*Z*) que são, ao mesmo tempo, produtos dos primeiros estágios e entradas para os estágios seguintes. No modelo proposto as organizações são avaliadas primeiramente por sua capacidade de utilização racional de recursos (Estágio 1: Produção) e posteriormente por sua capacidade de fornecer cursos ao maior número de servidores (Estágio 2: Entrega). Nesse sentido, uma unidade de treinamento eficiente no processo como um todo deve prover a maior quantidade de ações educacionais (*Z*) para a maior quantidade de indivíduos (*Y*) pelo menor custo (*X*). A seguir, detalham-se as implicações da definição do modelo de análise como um DEA em rede, assim como as demais extensões utilizadas.

3.7.1 Extensões ao modelo DEA adotadas

Conforme mencionado, a primeira extensão adotada para a melhor estimativa da eficiência nas organizações pesquisadas foi a modelagem de um modelo DEA em rede. O *Network DEA* caracteriza-se pela representação de um processo produtivo por meio de uma série de subprocessos interligados (Kao, 2017). A eficiência de cada etapa é analisada separadamente e posteriormente cria-se um índice agregado global do processo como um todo

por meio da multiplicação dos índices de eficiência das duas etapas. Assim, uma organização deve ser eficiente nas duas etapas para ser eficiente no processo. Essa é uma das técnicas básicas para criação de um índice agregado em um *Network* DEA, sendo que existem outras possíveis técnicas que consideram outros pressupostos e condicionantes para otimização da produtividade. O presente caso utilizará o modelo multiplicativo por ser o mais amplamente adotado e com pressupostos básicos para processos produtivos simples e com poucas variáveis (Cook, Zhu, Bi & Yang, 2010).

Além dessa extensão ao modelo básico, foi utilizada a análise intertemporal (Ferreira e Gomes, 2020), com abrangência do intervalo de cinco anos entre os anos de 2015 e 2019. Sob essa perspectiva, cada organização é considerada um caso independente para cada período de análise o que possibilita a constituição de uma amostra relevante para a distinção ao longo dos cinco anos de quais as provisões mais eficientes de treinamento (Ferreira & Gomes, 2020). Adicionalmente, buscou-se a realização de análises complementares voltadas a redução do viés determinístico do DEA, de forma que foram utilizadas análises de folgas (*slacks*) e *benchmarks*, a análise de índices de supereficiência para identificação de casos extremos e a utilização do *bootstrapping* para delineamento de intervalos de confiança para os índices de eficiência encontrados (Afonso, 2007; Cooper, Seiford & Tone, 2007; Ferreira & Gomes, 2020; Kao, 2017). Por fim, conforme Ferreira & Gomes (2020), buscou-se a consideração de influências de fatores contextuais, não integrantes do processo produtivo, concebidos como variáveis não-discrecionárias, sobre as quais assume-se a impossibilidade de manipulação por parte da DMU, mas que podem afetar a sua produtividade (Boueri et al., 2015).

Uma vez definidos os modelos e abordagens para alcance dos objetivos de pesquisa foi possível definir o *software* mais adequado para a análise. As etapas de coleta, tratamento e organização dos dados tiveram o apoio do *Excel*, um *software* proprietário de trabalho com planilhas eletrônicas. Ainda que seja possível a realização de rotinas básicas de DEA nesse *software* (Ferreira & Gomes, 2020), tendo em vista os objetivos de análise buscou-se por *softwares* específicos que permitissem a implementação dos modelos de análise necessários para a pesquisa. Assim, para a presente pesquisa, optou-se pela utilização do *software* aberto de análise estatística R. Dentre as principais vantagens do R, encontra-se o fato de ser um *software* livre, gratuito e colaborativo, congregando múltiplas ferramentas produzidas por diversos pesquisadores (Boueri et al., 2015). Dentre os pacotes para a realização do DEA encontra-se o pacote "*Benchmarking*", criado pelos pesquisadores Bogetoft & Otto (2019) que implementa uma variedade de métodos para estimar medidas de eficiências e comparação de tecnologias produtivas, dentre eles o DEA com suporte a variados pressupostos, variações e

modelos de análise (Bogetoft & Otto, 2019). Dessa forma, após estudo do referido pacote e realização de diferentes testes de modelos e variáveis para a análise, foi elaborado um *script* final no *software* R em que são detalhados os comandos e a sequência de análises realizadas na amostra. Os *outputs* do *script* foram manipulados e armazenados com apoio do *software* Excel. Um primeiro *script* detalha a análise da eficiência em todas as suas etapas (Apêndice G) e um segundo *script* dedica-se à análise das variáveis não-discrecionárias (Apêndice H). As seções seguintes detalham como foram testadas e selecionadas as variáveis para a análise e os fundamentos teóricos das rotinas de análise implementadas.

3.7.2 Teste de modelos e variáveis selecionadas

A coleta de dados forneceu uma série de informações acerca do processo de treinamento realizado nas organizações públicas federais. A partir da proposição do modelo em dois estágios foi definido um modelo inicial (Figura 26) o qual, após a realização de análises e testes foi redimensionado para o modelo apresentado ao final da seção.

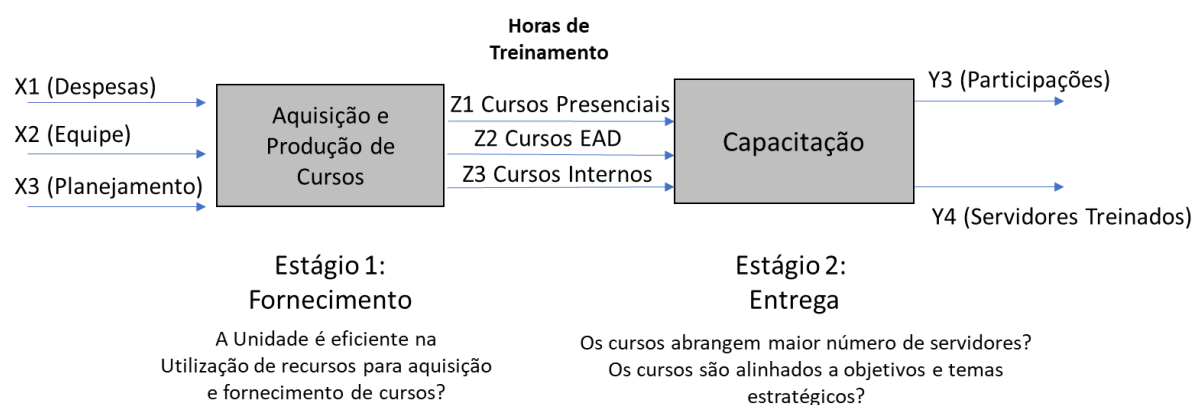


Figura 26. Modelo inicial

Fonte: Elaboração própria.

3.7.2.1 Variáveis de insumo (X)

3.7.2.1.1 Despesas com capacitação (X1)

Para análise da eficiência na aquisição e produção de ações educacionais (Estágio 1) se considerou como insumo as despesas totais com capacitação, definidas como a soma dos valores gastos com ações educacionais no período de um ano, envolvendo o custeio de inscrições, o pagamento de professores e a aquisição de recursos instrucionais. Não foram incluídos gastos com viagens e estadias, previstos separadamente em algumas bases de dados coletadas, em razão, primeiramente, de constituírem um objeto distinto do treinamento e pela ausência de padrão nesses tipos de registros.

Comparativamente à literatura, fatores financeiros como despesas e orçamento são frequentemente operacionalizados em estudos no campo educacional – 50% dos artigos da revisão realizada por De Witte et al. (2017) – e no âmbito corporativo, haja vista que 30% dos artigos encontrados na revisão de literatura desta pesquisa utilizaram recursos financeiros como insumos em suas análises. Na literatura, o termo “Orçamento” (*budget*) é mais utilizado, porém, este difere conceitualmente do termo “Despesas” adotado nesta pesquisa. “Orçamento” refere-se a uma medida acerca dos recursos reservados *a priori* para o treinamento, enquanto o termo “Despesas” é uma medida definida *a posteriori* como resultado da execução de uma determinada estratégia de provisão dos treinamentos, sendo mais condizente com o conceito de eficiência de gastos públicos que norteia a pesquisa, assim como a perspectiva somativa da avaliação de eficiência proposta.

A operacionalização dessa variável se deu por meio de medidas de valor financeiro referidas em produções acadêmicas como “Gasto anual com treinamento” e “Investimento em Treinamento”, sendo que a literatura aponta evidências de seu impacto em medidas de resultado financeiro (Kumar, 2019; Mehra et al., 2014). Sob o ponto de vista da literatura nacional, os recursos financeiros são considerados fatores relevantes para a efetiva implementação de políticas de capacitação (Camões & Meneses, 2016; Côrtes & Meneses, 2019).

3.7.2.1.2 Quantidade de integrantes da equipe da unidade de treinamento (X2)

A consideração do quantitativo de pessoas alocadas na unidade de treinamento tem como fundamento a previsão de fontes humanas como recursos a serem considerados na avaliação da eficiência em TD&E (Abbad et al., 2020). A revisão de De Witte et al. (2017) demonstra que o campo educacional operacionaliza variáveis similares, como “Pessoal” (*Personnel*), “Administradores” ou “Equipe de Apoio” (*Supporting staff*), como insumos em cerca de 30% dos estudos revisado nesse campo. Tal elemento foi também observado na revisão sistemática de literatura sobre eficiência de treinamento realizada em estudos acerca de organizações educacionais como o de Nojavan, Heidari & Mohammaditabar (2020) – “Quantitativo de Equipe” (*Number of Staff*), Bogetoft & Wittrup (2017) - “Horas de trabalho da equipe” (*Full Time Equivalent*). Já no âmbito corporativo, Kao et al. (2014) operacionaliza a variável “Equipe” (*Staff*) como sendo o número total de funcionários da equipe administrativa da instituição de treinamento.

A pesquisa preliminar verificou que alguns gestores entrevistados mencionaram esse fator como uma contingência para o desempenho de unidades de treinamento. Com efeito, Côrtes & Meneses (2019) destacaram em suas pesquisas a escassez de servidores nas unidades

de gestão de pessoas da administração pública federal como um fator de relevância. Dessa forma, essa variável foi operacionalizada em termos quantitativos representando o total de indivíduos, com quaisquer tipos de vínculo, que compõem a equipe responsável pela provisão de treinamentos na organização. Cabe mencionar que foram testadas operacionalizações com valores absolutos e relativos (Equipe/Força de Trabalho), porém não se considerou efetiva a segunda forma de operacionalização por estar relacionada a fatores não-discrecionários de ordem macro organizacional (Porte da organização) e que não eram determinantes da capacidade produtiva da equipe. Por fim, medidas qualitativas para mensurar a formação da equipe tiveram como obstáculo a insuficiência dos dados prestados pelas organizações em respostas aos instrumentos de pesquisa.

3.7.2.1.3 Planejamento da Capacitação (X3) como variável não-discrecionária

A especificação de uma variável relativa ao planejamento da capacitação visava a operacionalização da quantidade de ações planejadas pelas unidades organizacionais como resultado dos seus processos de ANT. No entanto, verificou-se que os Planos de Capacitação não eram homogêneos com relação a essa especificação, em que muitos especificavam temas gerais para o fornecimento de ações educacionais. Como alternativa, buscou-se a inclusão de variáveis categóricas dicotômicas representando a presença ou ausência de determinada atividade ou a quantidade de atividades da etapa de pré-treinamento (Salas et al., 2012) realizadas pelas organizações (e.g. ANT, Planejamento, Mapeamento de Competências). Apesar do DEA permitir a inclusão de variáveis dessas naturezas em suas análises, verificou-se que a quantificação dessa variável – de maneira dicotômica apenas – não geraria uma representação precisa de como as organizações realizam essas atividades, podendo levar a comparações de processos completamente distintos nas organizações. Além disso, ao se testar alguns modelos, organizações que não utilizam tais práticas podem figurar como sendo mais eficientes por serem capazes de produzir resultados mesmo com a ausência da ANT, do Mapeamento de Competências ou dos Planos de Capacitação, o que levaria à recomendação da supressão dessas atividades em outras organizações e não permitiria uma explicação adequada para a eficiência devido à falta de mais informações sobre a variável.

Ao se consultar a literatura, verifica-se que essa variável em específico é encontrada apenas em estudos qualitativos ou de estatística descritiva, em análises dos processos produtivos de unidades de treinamento, mas não são operacionalizadas em modelos DEA. Nesse sentido, compreendeu-se que essas variáveis poderiam ser consideradas variáveis exógenas, ou seja, um efeito prévio que interfere nas demais atividades de treinamento. Tal

abordagem foi utilizada, portanto, também para as demais variáveis coletadas por meio do formulário de processos de treinamento para fins de análise de suas possíveis associações com variações nos níveis de eficiência. Não obstante, estudos futuros deveriam buscar operacionalizar as etapas de integração mais precisa do sistema de TD&E com os insumos de informação oriundos dos processos de avaliação de necessidades e planejamento.

3.7.2.2 Variáveis intermediárias (Z)

3.7.2.2.1 Treinamentos (Z) – Carga Horária

Conforme mencionado, o Estágio 1 do modelo corresponde à eficiência na aquisição e produção de ações educacionais. De acordo com a literatura de Avaliação e o modelo teórico especificado para avaliação da eficiência em TD&E, atribuiu-se como uma variável de produto os quantitativos de ações educacionais implementadas pelas unidades de treinamento.

Dessa forma, o produto especificado para o primeiro estágio do modelo correspondeu aos cursos ofertados à clientela por meio da utilização dos recursos disponíveis. Para definição de sua operacionalização foram realizados testes com a quantidade de cursos ofertados e o produto entre carga horária e participantes. Também se buscou a estratificação do produto por modalidade educacional, porém observou-se os efeitos, previsto por Ferreira & Gomes (2020), de perda da capacidade de estimativa do modelo. Sendo assim, adotou-se apenas uma única variável, relativa à carga horária total dos treinamentos realizado no ano, sendo os aspectos instrucionais mencionados analisados qualitativamente.

Além disso, tendo em vista o conceito de treinamento, assim como os pressupostos de homogeneidade do DEA e sua sensibilidade a valores extremos, optou-se por excluir ações compreendidas como processos de “Educação”, ou seja, cursos de graduação, pós-graduação e similares, tais como bolsas para cursos de idiomas de longa duração. Primeiramente, porque tais ações possuem carga horária elevada e se estendem ao longo de múltiplos períodos de análise. Em segundo lugar, porque possuem objetivos distintos das ações de treinamento tradicionais, sendo mais associadas ao desenvolvimento pessoal dos servidores. Por fim, são implementadas de forma bastante diversa entre as organizações, ao contrário dos treinamentos, que são mais homogêneos em relação a incidência, formato e duração nas organizações pesquisadas.

A seleção dessa variável como um produto intermediário torna-a, simultaneamente, um insumo para a etapa de entrega do treinamento. Tal perspectiva é corroborada pela literatura, na qual a variável “Horas de Ensino” (*Teaching Hours*) é operacionalizada como insumo, normalmente para resultados de aprendizagem, em estudos no campo educacional (De Witte et

al., 2017). Com relação à revisão de literatura realizada nesta pesquisa, essa mesma variável é operacionalizada como horas de ensino por Salcedo (2020) e Bogetoft & Wittrup (2017) e como “Horas de Treinamento” por Kao et al. (2014) em estudos no ambiente corporativo, visando a análise do seu impacto nos resultados organizacionais.

3.7.2.3 *Variáveis de produto final (Y)*

3.7.2.3.1 Participações em treinamento (Y1) e quantidade de pessoal treinado (Y2)

Para fins dessa pesquisa, o produto final do processo de treinamento é o indivíduo capacitado. Idealmente, avaliações de treinamento devem buscar abranger a produção de resultados nos níveis dos indivíduos e da organização, que consistem nos seus principais objetivos. No entanto, considera-se que instrumentos de avaliação específicos devem ser aplicados por meio de procedimentos formais e homogêneos em todas as organizações para que as medidas relativas aos resultados de TD&E sejam analisadas. Logo, o modelo proposto se limita a uma avaliação de eficiência em perspectiva estritamente operacional para decisões de implementação de treinamentos.

Para operacionalização dessa variável foram testadas duas abordagens: a quantidade de vagas ofertadas e a quantidade de indivíduos únicos capacitados. Os bancos de dados fornecidos pelas organizações participantes permitiam a contabilização de quantos indivíduos da força de trabalho haviam sido treinados em cada ano, sendo possível inclusive a criação de um índice percentual de proporção da força de trabalho capacitada. Tal perspectiva distingue-se da contabilização de Participações que permite que uma mesma pessoa seja contada mais de uma vez ao participar de múltiplos treinamentos. A quantificação de participações, no entanto, era menos afetada pelo porte da organização – organizações com poucos funcionários sempre capacitariam poucos indivíduos – e a ponderação por meio de um índice percentual criava um limite artificial (100% da força de trabalho capacitada) para a variável que influenciava a adequada mensuração de produtividade.

Por esse motivo, adotou-se como medida a quantidade de participações ou vagas totais de treinamento oferecidas no ano. Conforme Ferreira & Gomes (2020), tais questões podem ser solucionadas com testes de correlação ou diagramas de dispersão para averiguar se as variáveis não são aspectos de um mesmo fenômeno. Nesse sentido, conforme verifica-se na Figura 27, este é o caso para ambas as variáveis, quantidade de participações e proporção de indivíduos treinados, posto que há indícios de correlação entre elas. De toda maneira, as vagas ofertadas representam um desempenho objetivo e factual das unidades de treinamento, enquanto a

proporção de indivíduos capacitados pode ser compreendida como um efeito ou um resultado almejado, mas não como um produto direto e gerenciável do processo de capacitação.

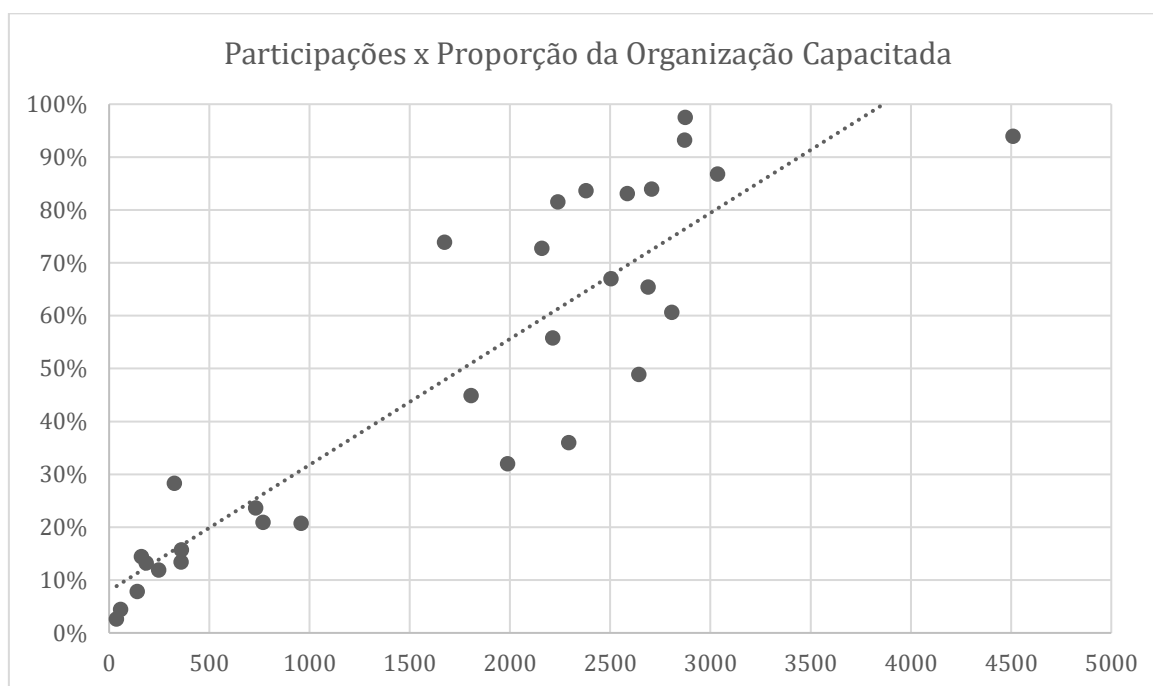


Figura 27. Quantidade de participações x percentual da força de trabalho capacitada

Nota. Fonte: Elaboração própria.

Com base nos dados da Figura 27, considera-se que quanto maior a quantidade de oportunidades de treinamento uma organização fornece, pode-se esperar que um maior percentual de sua força de trabalho será capacitado. Nesse sentido, considerou-se que a variável selecionada permite representar um critério de equidade para o desempenho do sistema de TD&E, o qual também possui relevância no campo de avaliação educacional (Johnes et al., 2017). Quando considerada essa literatura, verifica-se que a quantidade de pessoas formadas é o principal produto na maior parte dos artigos do campo educacional e a maneira mais comum de se operacionalizar esse resultado (De Witte et al., 2017). Em contraste, Cubbin & Zamani (1996) adotaram perspectiva similar à presente pesquisa, no âmbito do treinamento no setor público, contabilizando a quantidade de certificados (i.e. participações) emitidos anualmente por instituições públicas de educação profissional.

3.7.2.4 Teste de Correlação de variáveis

Como segunda etapa do processo de seleção de variáveis para definição de um modelo DEA, deve-se proceder à sua análise estatística. Para tanto, procedeu-se conforme sugerido por Ferreira e Gomes (2020), a um teste de intercorrelação entre as variáveis (Tabela 14).

Tabela 14
Teste de Correlação de variáveis

	X1	X2	Z	Y1	Y2
X1	1	0,34	0,76	0,81	0,77
X2		1	0,49	0,53	0,63
Z			1	0,71	0,81
Y1				1	0,90
Y2					1

Nota. Fonte: *Elaboração própria.*

Conforme Ferreira & Gomes (2020), é esperado que haja níveis moderados de correlação entre as variáveis de um modelo DEA de forma que aumentos em variáveis de insumo provoquem aumentos esperados em variáveis de produto. Por outro lado, caso variáveis da mesma categoria tenham uma correlação muito alta (acima de 0,80), elas são provavelmente substitutas e uma delas pode ser eliminada. Seguindo-se tais recomendações verifica-se que as variáveis de insumo Equipe (X1) e Despesas (X2) estão adequadamente correlacionadas com a variável horas de treinamento (Z). Esta por sua vez correlaciona-se à quantidade de participações (Y1). Como já deliberado anteriormente, confirmou-se, com essa análise, que a variável relativa à proporção da força de trabalho capacitada (Y2) pode ser eliminada, pois correlaciona-se em 90% com a variável Y1 relativa à quantidade de participações, sendo a única necessária para integrar o produto final.

3.7.2.5 Variáveis ambientais exógenas e não-discrecionárias

Após o cálculo dos índices de eficiência de uma determinada amostra, é possível investigar os efeitos de variáveis não-discrecionárias, exógenas ou fixas que se caracterizam por não serem imediatamente ajustáveis ou gerenciáveis pelas DMUs (Ferreira & Gomes, 2020). A análise proposta tem caráter exploratório, em razão de não haver literatura que especifique variáveis interferentes na eficiência de treinamento. Assim, foram realizadas análises acerca da influência de variáveis, consideradas não-discrecionárias, categóricas e dicotômicas, relativas à presença e ausência dos seguintes processos e práticas conforme relatos das organizações pesquisadas:

- Avaliação de Necessidades de Treinamento;
- Mapeamento de Competências;
- Planos Anuais de Capacitação;
- Avaliação de Reação;
- Avaliação de Aprendizagem;
- Avaliação de Impacto;
- Planejamento Estratégico com objetivos de capacitação; e

- Sistema informatizado de capacitação.

Além desses aspectos, por meio de análises documentais, foram levantados outros fatores passíveis de análise como variáveis não-discrecionárias que visam captar aspectos da estrutura organizacional (Camões & Meneses, 2016; Côrtes & Meneses, 2019). Sendo assim, foram definidas três variáveis de interesse para a pesquisa, as quais foram dicotomizadas para fins de comparação de grupos.

- **Porte Organizacional:** relativo ao quantitativo total de servidores na organização, dicotomizada entre organizações com mais (1) e menos (0) de 1.100 servidores (Mediana da amostra).
- **Número de unidades descentralizadas:** relativo à distribuição da organização em unidades da federação, sendo as organizações classificadas naquelas que possuem múltiplas unidades administrativas no território nacional (1) e aquelas que possuem a maior parte dos servidores atuando em uma única localização principal, sede da organização (0).
- **Níveis hierárquicos:** Relativo à quantidade de instâncias decisórias acima da unidade de treinamento, refletindo sua autonomia estrutural (Fonseca et al., 2013). As organizações foram classificadas entre aquelas em que a unidade de treinamento está a até três níveis da cúpula organizacional (1) (e.g. Presidência, Diretoria de Gestão de Pessoas, Unidade de Treinamento) e aquelas que estão inseridas em estrutura mais abaixo na hierarquia organizacional (0).

3.7.3 Modelo e Base de Dados Final

Após a realização dos testes e definições supracitados, o modelo final de análise consistiu em um total de quatro variáveis, tendo apresentado boa capacidade de distinção das DMUs com relação aos índices de eficiência e respectivos fatores de produção. A base de dados quantitativos utilizados na análise pode ser encontrada no Apêndice E enquanto os dados com os fatores não discrecionários dicotomizados pode ser encontrada no Apêndice F. A Figura 28 apresenta o modelo final proposto para avaliação da eficiência das unidades de treinamento na Administração Pública Federal.

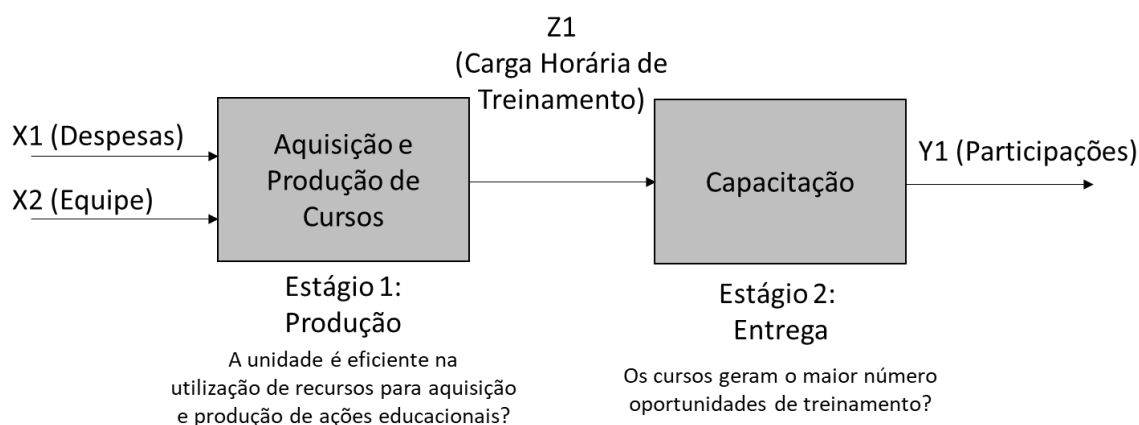


Figura 28. Modelo Final de Análise

Nota. Fonte: Elaboração própria.

3.8 Procedimentos e Técnicas de Análise

A análise dos dados foi realizada em múltiplas etapas no *software* estatístico R com o pacote *benchmarking*. Partiu-se da análise de eficiência com os modelos básicos da literatura: CCR e BCC (Ferreira & Gomes, 2020) e procedeu-se à sua complementação com técnicas que visaram ampliar a compreensão do fenômeno e reduzir possíveis vieses determinísticos do DEA, assim como possíveis limitações devidas ao tamanho da amostra. As subseções a seguir detalham as abordagens, métodos e rotinas empregados para análise da eficiência e seus fatores.

3.8.1 Análise Intertemporal de dados em Painel

Como primeiro aspecto, denominam-se dados em painel uma combinação de dados organizados em uma série temporal e que permite fazer uma análise da eficiência para todo o período (intertemporal), mas levando em consideração os desempenhos históricos da série (Ferreira & Gomes, 2020). Para a presente pesquisa, optou-se pela realização da análise intertemporal, uma abordagem na qual se constrói uma única fronteira de eficiência com cada caso em cada período sendo considerado uma DMU independente e não se analisando progressos tecnológicos no conjunto de organizações. Métodos mais avançados, mas que dependem de uma amostra maior, tais como a Análise de Janelas e o Índice de Malmquist permitem avaliar o progresso das DMUs ao longo do tempo e comparar curvas de eficiência de diferentes períodos, mensurando as variações da tecnologia produtiva ao longo do tempo (Ferreira & Gomes, 2020).

3.8.2 *Análise da eficiência e definição de pressupostos de análise*

A análise da eficiência pode ser realizada com base em dois métodos básicos: o método CCR (Charnes et al, 1978) e o método BCC (Banker et al., 1984). O primeiro pressupõe uma relação linear de retornos constantes entre insumos e produtos, enquanto o segundo admite uma relação não linear com retornos variáveis, que podem ser crescentes ou decrescentes. Tais distinções podem determinar como será traçada a fronteira de eficiência e como a escala ou porte organizacional influenciará os resultados, o que modifica, conseqüentemente, quais organizações serão consideradas ineficientes (Rosano-Peña, 2008; Ferreira & Gomes, 2020; Kao, 2017).

Nesse sentido, para averiguar qual dos pressupostos melhor se adequa à amostra, pode-se realizar a análise da eficiência por ambos os modelos e, posteriormente, realizar um teste de hipóteses para averiguar diferenças de médias ou distribuições probabilísticas para os resultados encontrados. Para a presente pesquisa optou-se pela utilização do teste não paramétrico denominado teste de Kolmogorov-Smirnov - KS, tendo em vista a sua maior flexibilidade e robustez para valores que podem não atender aos requisitos dos testes paramétricos e por se apresentar como um método adequado a pequenas amostras (Firmino, 2015).

O Teste KS normalmente é utilizado em testes de normalidade, para comparação de uma distribuição empírica com uma distribuição normal. No presente caso, a comparação será feita entre distribuições de resultados de índices de eficiência CCR e BCC, visando atestar se elas são distintas. O teste é realizado, portanto, com base nas seguintes hipóteses:

H₀: Não há diferença entre os resultados.

H₁: Há diferença entre os resultados.

O teste realizado é bilateral com nível de significância de 0,05 (Firmino, 2015). Caso a hipótese nula não seja rejeitada, não há distinção entre os índices de eficiência, logo, pode-se adotar o modelo CCR. Por outro lado, quando a hipótese nula é rejeitada, há diferenças entre as amostras devido à convexidade da curva de eficiência da relação insumo-produto do modelo BCC, indicando que a utilização desse modelo é um melhor ajuste à amostra. Deve-se ter em consideração, no entanto, que a convexidade do BCC reduz o impacto decorrente de escalas de produção ineficientes. Caso se deseje analisar essa ineficiência deve-se obter um índice baseado na razão entre os índices de eficiência CCR e BCC, sendo esse procedimento o recomendado

por Ferreira & Gomes (2020) para que se possa distinguir ineficiência técnica da ineficiência de escala.

Um segundo aspecto que influencia na determinação dos índices de eficiência é a orientação do modelo de análise. A orientação para insumos privilegia a economia desses recursos, enquanto a orientação para produtos privilegia a maior produtividade. Para a presente pesquisa considera-se que as unidades de treinamento são incentivadas a utilizar todo o orçamento disponível na aquisição e produção de unidades de treinamento (Ribeiro, 2010; Fonseca et al. 2017) e que devem buscar estender a oferta de cursos ao máximo de indivíduos na organização (Brasil, 2019a). Conclui-se, portanto, que a orientação para o produto é a mais indicada e coerente com o contexto de análise.

Tendo em vista tais pressupostos e testes realizados, o modelo para análise da eficiência será o de retornos variáveis (*Variable Returns to Scale – VRS*) com orientação para produto (*output*)⁸. A análise é realizada para cada um dos estágios do *Network-DEA* e a obtenção de um valor global de eficiência é realizada por meio da multiplicação dos índices de cada etapa (Cook et al., 2010). Para melhor interpretação dos índices de eficiência orientados para produto poderão ser consideradas as razões inversas dos resultados para que se padronize a análise com índices que variam entre 0 e 1, facilitando a compreensão. Além disso, complementarmente serão consideradas possíveis ineficiências de escala na análise (Ferreira & Gomes, 2020), sobretudo no Estágio 2 em que o Teste KS entre as eficiências CCR e BCC resultou significativo.

3.8.3 *Outliers e Supereficiência*

A obtenção de índices de eficiência e a consequente ordenação de organizações com relação ao seu desempenho pode ser fortemente influenciada por valores extremos e vieses amostrais (Rosano-Peña, 2008). Para o primeiro caso, um dos possíveis remédios é a utilização de rotinas para detecção de *outliers*, casos com valores extremos que podem distorcer a fronteira de eficiência (Rosano-Peña, 2008). O pacote “*Benchmarking*” possui uma rotina específica para detecção de possíveis *outliers*. Se os respectivos casos se constituírem como firmas eficientes, uma alternativa é a realização da análise de supereficiência que consiste na retirada da restrição matemática ao índice, permitindo que os índices de eficiência variem de 0 a infinito (Ferreira

⁸ Definição baseada nas estatísticas do teste para KS para eficiências CCR e BCC com orientação para produto.

Estágio 1: $D = 0.267$, $p\text{-valor} = 0.2365$, não significante, logo não há diferença entre os resultados.

Estágio 2: $D = 0.56667$, $p\text{-value} = 0.000131$, significante, logo há diferenças e indicada a adoção do BCC (VRS).

& Gomes, 2020). Sendo esses últimos casos encontrados, pode ser indicada a sua remoção da análise. A análise preliminar da amostra demonstrou a presença de um *outlier* de relevância no primeiro estágio (D17), o qual apresentou valores extremos na variável relativa às despesas de capacitação. No entanto, optou-se por sua não remoção devido ao julgamento de que este tipo de valor extremo é relevante para o objetivo de pesquisa, além de que o caso não apresentou supereficiência (0,93).

3.8.4 DEA Bootstrap

Uma outra abordagem para redução de vieses determinísticos é a realização de técnicas de *bootstrapping* para cálculo dos índices de eficiência (Bogetoft & Otto, 2011). Sumariamente, o *bootstrapping* é uma técnica estatística que simula uma mesma análise de uma determinada amostra múltiplas vezes e fornece um intervalo de confiança com os possíveis resultados que a análise pode apresentar, caso sejam admitidas variações na sua distribuição. Para isso a técnica simula diferentes versões do mesmo conjunto de dados “n” vezes, por meio de um processo de reamostragem aleatória livre que termina por simular a análise em uma população maior de casos quão maior seja o número “n” de testes⁹ (Bogetoft & Otto, 2011). Tal rotina existente no pacote “*Benchmarking*” foi utilizada com o objetivo de criar intervalos de confiança para a melhor avaliação dos índices de eficiência das organizações analisadas e, se possível, evitar a remoção de eventuais casos extremos.

3.8.5 Modelos aditivos, análise de folgas (slacks) e benchmarks

As perspectivas clássicas de eficiência assumem que as organizações localizadas na fronteira eficiente se encontram em um ponto ótimo de produtividade e relação insumos-produtos (Ferreira & Gomes, 2020). Além disso, as organizações ineficientes devem ajustar seus insumos ou produtos, mantendo as demais variáveis fixas, para alcançar a eficiência. Uma perspectiva alternativa, chamada de modelo aditivo, pode ser assumida para interpretação de ambos os casos visando a proposição de ajustes visando a eficiência (Cooper, Seiford & Tone, 2007; Ferreira & Gomes, 2020).

Conforme a Figura 29, o modelo aditivo propõe, primeiramente, que DMUs ineficientes, no caso representadas pela DMU “D”, podem utilizar uma combinação de ajustes de insumos e produtos visando percorrer uma menor distância até a fronteira eficiente (Cooper, Seiford & Tone, 2007; Ferreira & Gomes, 2020).

⁹ Nessa pesquisa adotou-se N = 10.000.

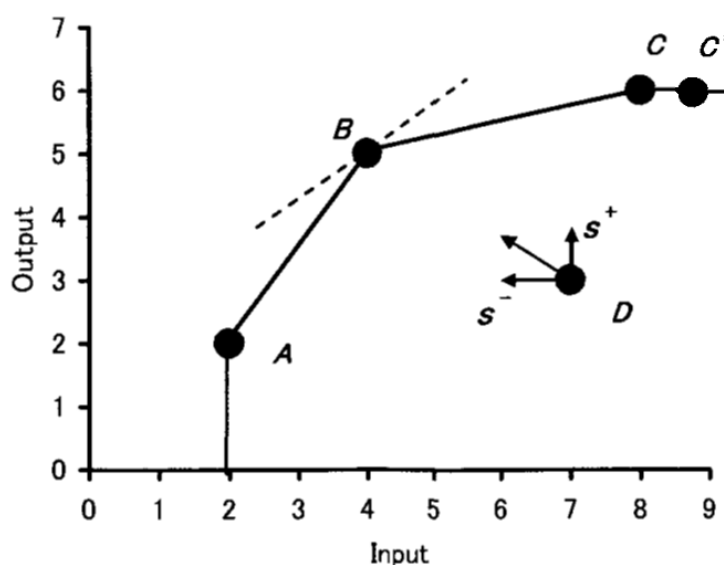


Figura 29. Modelo Aditivo

Nota. Fonte: Adaptado de Cooper, Seinfeld & Tone (2007).

O modelo aditivo busca realizar o cálculo de uma função de otimização que melhor se ajuste a cada DMU ineficiente para levá-la à eficiência. Com isso, são indicados os quantitativos que devem ser ajustados em cada insumo ou produto. A indicação de ajuste se baseia no cálculo da distância em relação a uma DMU eficiente, a qual torna-se, portanto, uma referência para a respectiva DMU ineficiente, sendo referida como *benchmark* ou parceira (*peers*). Assim, um índice “ λ ” (lambda) de 0 a 1, indica o quanto uma determinada DMU eficiente serve de referência para uma determinada DMU ineficiente. Quanto mais índices “ λ ” uma DMU eficiente possui, mais importante ela é em toda a amostra (Ferreira & Gomes, 2020).

Por fim, uma extensão dos modelos aditivos para cálculo da eficiência permite a identificação da chamada eficiência “fraca”, a qual está representada pelo ponto C’. Muito embora esse ponto C’ esteja sobre a fronteira de eficiência, pode-se observar que sua distância em relação ao ponto C implica na possibilidade de redução dos insumos, sem redução dos produtos. Tais insumos são considerados desperdícios, ou folgas (*slacks*) e são indicadores de que mesmo uma unidade eficiente pode ser capaz de realizar ajustes para se tornar verdadeiramente eficiente (Ferreira & Gomes, 2020).

Tais rotinas de análise estão presentes no pacote “*Benchmarking*” e serão utilizadas na análise visando a compreensão de como as unidades de treinamento podem ajustar suas variáveis para obtenção de maior eficiência e, sobretudo, para identificar quais unidades e respectivas estratégias são consideradas os *benchmarks* mais relevantes na amostra.

3.8.6 *Análise de variáveis não-discricionárias: Modelo de dois estágios.*

A última etapa de análise realizada consistiu no exame das variáveis não-discricionárias anteriormente definidas a partir dos dados coletados junto às organizações pesquisadas. Primeiramente, deve-se ter em mente que a análise de fatores é considerada uma etapa posterior à análise da eficiência, sendo esse modelo chamado de modelo de dois estágios, o qual não se confunde com os estágios do DEA em rede. Nessa acepção, o primeiro estágio dedica-se ao cálculo da eficiência e o segundo à estimação do efeito que variáveis não-discricionárias possam ter sobre os valores observados (Boueri et al., 2015).

Segundo Boueri et al. (2015) e Afonso (2007), uma abordagem comum para análise de variáveis exógenas ou não-discricionárias é a realização de uma regressão linear Tobit para estipular os seus efeitos nos índices de eficiência. Por outro lado, Boueri et al. (2015) e Ferreira & Gomes (2020) apresentam um método simplificado de análise, baseado inteiramente em DEA, em que não é necessária a previsão de hipóteses ou a compreensão prévia acerca dos fatores, bastando apenas a subdivisão da amostra em subgrupos para comparação dos seus valores de eficiência. Tais grupos podem ser por exemplo organizações públicas e privadas (Boueri et al., 2015). Para a presente pesquisa, optou-se pela utilização de valores dicotômicos para operacionalização das variáveis ambientais, tornando possível a definição de dois subgrupos requeridos para esse tipo de análise.

Seguindo-se roteiro proposto por Boueri et al. (2015) com base em Charnes, Cooper e Rhodes (1981), foram calculadas as eficiências para cada grupo separadamente e então projetadas todas as unidades para a respectiva fronteira de eficiência do grupo – no caso, multiplicando os fatores de produção pelo índice de ineficiência correspondente, tornando-as eficientes. Em seguida, foi rodado o DEA com as amostras de ambos os grupos misturadas, de forma que todas são comparadas como se estivessem operando com eficiência máxima. Por fim, com utilização de um teste comparativo – no presente caso, o Teste KS (Firmino, 2015) – foi aferida a significância da diferença entre os índices de eficiência de cada grupo para estabelecer se a presença ou ausência do respectivo fator favoreceu ou não o desempenho de um dos grupos.

3.8.7 *Interpretação dos índices de eficiência*

A interpretação dos dados oriundos da aplicação do DEA deve ser cautelosa e se restringir àquilo que os dados permitem inferir. Para o presente caso, a avaliação da eficiência das unidades de treinamento foi realizada após múltiplas rodadas de análise para definição de um modelo capaz de se adequar à amostra de organizações, aos dados disponíveis e aos

objetivos de pesquisa. Adicionalmente, variadas extensões ao modelo foram aplicadas para mitigação de eventuais vieses determinísticos.

Os pressupostos de análise mais ajustados aos dados assumem uma função produção do treinamento com retornos variáveis à escala (*variable returns to scale* – VRS). Além disso, adotou-se a orientação para o produto (*output*), que visa otimizar a eficiência das unidades por meio do incremento dos produtos, mantendo-se a mesma quantidade de insumos. Sob uma perspectiva de orientação aos produtos, os índices de eficiência são iguais ou superiores a “1,0”, sendo maiores quanto maior for a proporção do incremento necessário aos produtos. Por exemplo, um coeficiente “1,0” é indicativo de uma produção perfeitamente eficiente, enquanto um coeficiente “1,25” indica que a unidade deve aumentar todos os bens e serviços produzidos em 25% para ser eficiente (Rosano-Peña, 2008).

Tendo em vista as técnicas de análise empregadas nessa pesquisa, foi considerada mais adequada a inversão dos escores de eficiência para que se ativessem a um escore entre 0 e 1,0, visando facilitar a interpretação e a manipulação dos dados. Dessa forma, o índice exemplificado de “1,25” será de $0,8 = 1/1,25$. Com isso a interpretação é de que a unidade está operando a apenas 80% da capacidade necessária para que se torne eficiente. Uma organização com um índice de “2,0” deveria dobrar a sua produção para se tornar eficiente, de forma que ela está, portanto, produzindo apenas metade, $50\% = 1/2$, do que deveria para ser eficiente, tendo como referência o desempenho das demais organizações da amostra.

Tal interpretação será adotada no decorrer da análise refletindo a capacidade das unidades de utilizar seus recursos para ofertar a maior quantidade de horas de treinamento (Estágio 1: Produção) para a maior quantidade de participantes (Estágio 2: Entrega). Vale lembrar que não estão abrangidas na análise ações relativas à educação (e.g. graduação, pós-graduação, etc.) e outras que venham a ter longa duração, visando à homogeneidade das ações de treinamento para comparação. Finalmente, diante de todo o exposto, a seção seguinte apresenta os resultados obtidos após a execução de todos os procedimentos de análise explicados.

4 Resultados e discussão

Esta seção destina-se a apresentar os resultados de pesquisa obtidos ao longo de quatro subseções, nas quais são detalhados os resultados referentes ao segundo objetivo específico do estudo de realizar a avaliação da eficiência, comparando-se as organizações estudadas, e ao terceiro objetivo de verificar as influências de fatores interferentes na eficiência.

A primeira subseção realiza uma avaliação conjunta das organizações, apresentando os valores médios da eficiência técnica de produção de cursos – Estágio 01 do modelo – que versa, principalmente, sobre a eficiência do gasto para capacitação; da eficiência técnica de entrega – Estágio 02 – que versa sobre a geração de oportunidades de capacitação promovendo seu maior alcance; e, por fim, da eficiência global, produto dos dois estágios. Para essa subseção, promove-se uma discussão visando compreender, interpretar e contextualizar os resultados gerais.

A segunda subseção analisa de forma mais aprofundada cada um dos estágios de eficiência a partir dos resultados da Análise Envoltória de Dados. Serão detalhados e discutidos os desempenhos das organizações em cada Estágio, apresentadas as respectivas fronteiras de eficiência e considerações sobre intervalos de confiança. As organizações são comparadas entre si e discute-se os principais ajustes relativos a insumos e produtos nas DMUs ineficientes, tendo em vista os *benchmarks*, ou as melhores práticas detectadas no âmbito da amostra.

A terceira subseção analisa a eficiência global das unidades e apresenta a análise de possíveis relações com determinadas características das unidades e programas ofertados de TD&E, tais como os quantitativos de equipe e de investimento, assim como as estratégias de capacitação com relação à modalidade (presencial ou à distância) e o tipo de provisão (curso externo ou interno). Também são apresentados os resultados da análise do impacto de fatores condicionantes na eficiência, sendo considerados processos de TD&E como a ANT e Mapeamento de Competências, assim como aspectos organizacionais como o alinhamento estratégico, porte e descentralização organizacional. Por fim, a última subseção estende a discussão dos resultados em perspectiva mais qualitativa, buscando analisar a eficiência de cada uma das organizações frente aos seus processos de TD&E e suas decisões acerca dos tipos de cursos ofertados.

A discussão dos resultados enseja cautela na sua interpretação, uma vez que, primeiramente, eles dirão respeito apenas às organizações da amostra, não cabendo extrapolações, generalizações ou comparações com outros casos não abrangidos pela pesquisa. O objetivo dos modelos DEA não se volta para a explicação e generalização de relações, mas

para determinação das relações insumo-produto entre um determinado conjunto de DMUs visando à sua avaliação. As associações entre eficiência e outras variáveis componentes das unidades e programas de treinamento são realizadas em termos de interpretações possíveis, mas não estão submetidas a testes estatísticos, com exceção das análises de fatores condicionantes, de forma que visam antes promover a discussão e suscitar novas questões do que estabelecer um diagnóstico definitivo acerca da dinâmica entre eficiência e treinamento. Não obstante, a seção demonstra que o DEA cumpre o objetivo de promover a análise e discussão de alternativas de investimento e implementação do TD&E que impactam na sua eficiência e estabelece uma perspectiva com vantagens distintas para a análise de programas de treinamento.

4.1 Avaliação conjunta da eficiência das unidades de treinamento

A avaliação da eficiência tem como um de seus objetivos verificar as alternativas de investimento no âmbito de programas organizacionais, não versando acerca do mérito e valor das ações implementadas, mas se houve uma alocação ótima dos recursos (Suchman, 1967). O modelo de análise adotado nesse estudo verifica a eficiência alocativa em dois estágios. O primeiro estágio refere-se à adequação entre os recursos utilizados e seus resultados na provisão da maior quantidade de ações de TD&E, em termos de horas de capacitação. O segundo verifica a adequação entre o tempo de capacitação e o quantitativo de participações de servidores nessas ações, de forma que o objetivo final do modelo é avaliar a capacidade operacional do sistema de capacitar o maior número de pessoas possível, haja vista os recursos empreendidos.

Para a avaliação conjunta da eficiência das unidades de treinamento, a Figura 30 apresenta o ponto de partida da análise, em que é possível verificar a evolução da média dos índices de eficiência das 06 organizações pesquisadas ao longo do período de 2015 a 2019, comparativamente ao total de investimentos em treinamento realizados (em milhões de R\$).

Conforme mencionado, o primeiro estágio da análise – “Produção”, em azul - representa a eficiência das organizações na aquisição e produção de treinamentos a partir dos seus recursos humanos (a variável de insumo X1, Equipe) e financeiros (a variável de insumo X2, Despesas). De forma geral, as organizações possuem uma eficiência média relativamente alta, variando entre 0,59 em 2018 e 0,80 em 2017. Com relação ao segundo estágio – “Entrega”, em laranja – verifica-se que as organizações são menos eficientes no que se refere à promoção do maior número de participações em treinamento, com um índice variando entre 0,42 em 2019 e 0,55 em 2018. Por fim, o escore global de eficiência, oriundo do produto dos dois escores anteriores, é relativamente baixo (em torno de 0,30) e ao mesmo tempo estável ao longo do período analisado.

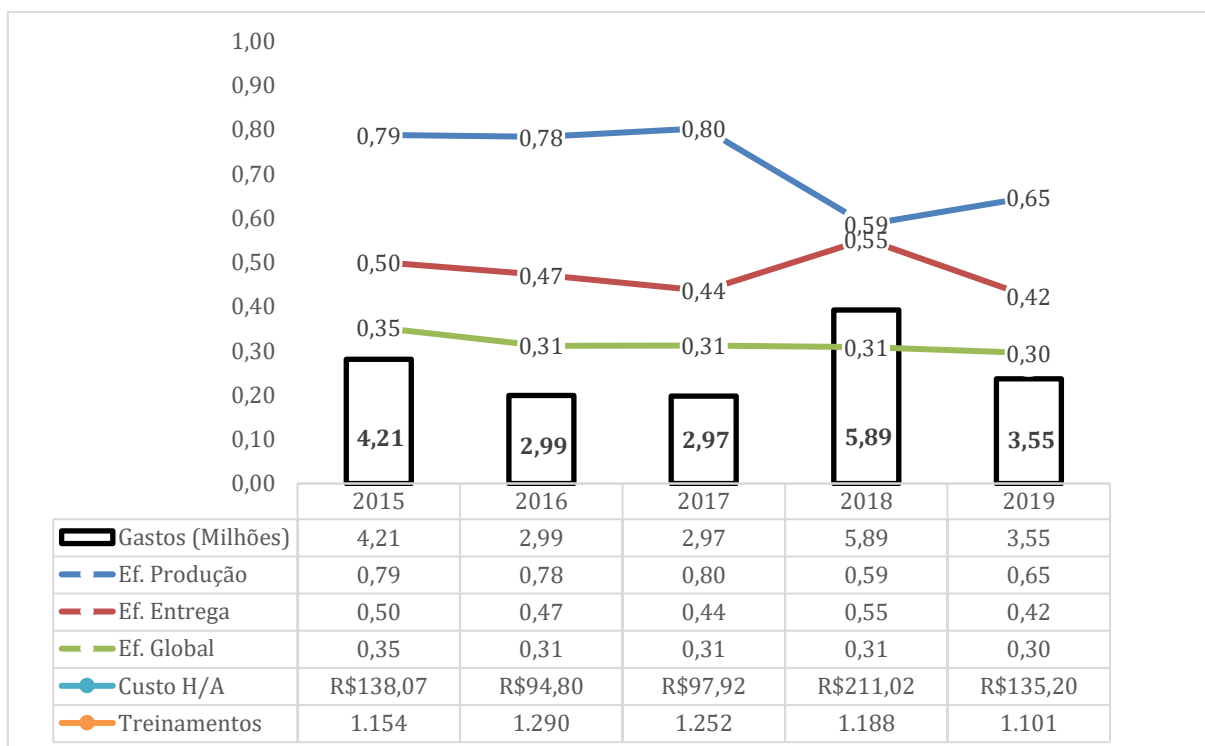


Figura 30. Evolução dos índices de eficiência por período

Fonte: Elaboração própria.

Avaliando-se apenas os índices médios de eficiência, estes sugerem a possibilidade de aperfeiçoamento dos processos de provisão de treinamento nas organizações, mas, sobretudo, no que se refere ao estágio 2: alcance de maior número de servidores. Dado o índice de eficiência de entrega em torno de 0,50, depreende-se que poderiam ser ofertados o dobro de oportunidades de treinamento. Os índices relativamente altos de eficiência de produção, sugerem, portanto, que as organizações possuem maior orientação para a eficiência na contratação de cursos, mas que isso não se reflete, necessariamente, no maior número de oportunidades de capacitação. Vale lembrar que o número de participações apresentou uma alta correlação (0,90) com o indicador da proporção de pessoas capacitadas da força de trabalho, de forma que se reflete, portanto, no critério da equidade no acesso à capacitação.

Com isso, pode haver um indicativo de que os cursos contratados, ainda que econômicos, não estejam cumprindo plenamente os objetivos do Sistema de TD&E, posto que tão importante quanto a eficiência de produção é a capacidade de entrega desses cursos para o maior número de pessoas, visando ampliar a equidade e potencializar os resultados da capacitação. Considerando a literatura de avaliação de programas, os resultados operacionais, tais como o alcance adequado do público-alvo, são um objetivo intermediário relevante para o alcance dos seus objetivos finais, relacionados à eficácia e efetividade (Suchman, 1967), sendo

também utilizado como medida de desempenho na literatura de avaliação educacional (De Witte & Lopes-Torres, 2017).

No entanto, parece haver certa concorrência entre esses objetivos de economia de recursos e maior alcance da capacitação, quando se observa principalmente o ano de 2018. Para a amostra estudada, quando alcançaram o maior índice de eficiência na entrega de cursos, as organizações o fizeram em prejuízo da eficiência produtiva. Por outro lado, a eficiência global foi mantida, o que implica que podem estar retratadas duas estratégias de provisão para o mesmo índice de eficiência global: uma focada em economia de custos e outra focada em maior número de participações. Considerados os indicadores de investimento, verifica-se que o enfoque nessa segunda estratégia ensejou mais custos, sendo 2018 o período em que houve maior investimento em treinamento (5,89 milhões). O número de treinamentos ofertados, no entanto, foi relativamente estável ao longo de todo o período (em média 1.197 cursos), ou seja, o maior alcance foi obtido com a oferta de cursos de maior valor em 2018.

A economia de gastos é um princípio relevante no setor público, porém essa deve ser equilibrada com a concretização de objetivos (Modesto, 2000). Sob um ponto de vista do processo de TD&E, podem estar presentes falhas de planejamento, em que, quando ausentes mecanismos eficazes de ANT e planejamento instrucional, a organização pode se tornar antes um centro de contratações de cursos de curta duração para atendimento de demandas pontuais (Meneses et al., 2010) do que uma unidade capaz de desenvolver cursos estratégicos e políticas eficazes que promovam maior adesão de servidores públicos e adicionem valor à organização (Borges-Andrade et al., 2012). Com efeito, como será discutido adiante, há uma prevalência de contratações de cursos externos de curta duração nas organizações da amostra.

Essa dificuldade de planejamento pode estar associada a questões do contexto normativo-institucional, em que as organizações públicas, em geral, não possuem discricionariedade acerca da quantidade de recursos financeiros disponíveis para capacitação, de forma que possuem um orçamento estipulado externamente que não considera os seus planos e estratégias de capacitação. Além disso, estão sujeitas a regras de contratação que priorizam a aquisição de produtos e serviços pelo menor valor de mercado (Ribeiro, 2010). Tais aspectos podem não ser adequados quando se trata da provisão de treinamentos, uma vez que esses devem ser planejados e ofertados antes pela sua capacidade de atender às necessidades de desenvolvimento dos servidores e aos objetivos organizacionais (Freitas & Mourão, 2012; Holton, 1996), e não apenas pelo seu custo (Abbad et al., 2020). Bach (2010), por exemplo, observa que há uma tendência no setor público de se considerar a gestão de pessoas como um

custo e não como um investimento, argumento também apresentado por Pires et al. (2009) e por Camões e Meneses (2016) para o caso da capacitação no setor público federal brasileiro.

Como evidência desses fatores contextuais, verificou-se nas entrevistas preliminares que as organizações estiveram sujeitas a fortes restrições orçamentárias nos anos de 2016 e 2017 decorrentes de crises fiscais, em que os recursos para capacitação foram suspensos no primeiro ano e reduzidos no segundo. Assim, se observa uma redução relevante nos indicadores de investimento em treinamentos neste período, conforme demonstrado na Figura 30. Levadas a reduzir o custo dos treinamentos, as organizações obtiveram maiores índices de eficiência de produção nestes anos, mas não de eficiência de entrega. Cabe mencionar também que em 2018, a mudança de governo refletiu-se nas organizações e, em alguns casos, nas unidades de treinamento, com mudanças de gestores responsáveis pelos departamentos de pessoal, o que pode ter ensejado em alterações em seus processos, planos e programas de treinamento.

Tais questões levantam, antes, um grande número de questões que podem ser melhor investigadas por estudos posteriores acerca da dinâmica e dos resultados dos investimentos em capacitação de larga escala no contexto do setor público federal. O escopo do presente estudo permite apenas o cotejamento das possíveis razões contextuais para o comportamento das organizações observadas. Estudos mais específicos podem, similarmente ao que se observa no campo da Educação (Johnes et al., 2017), estudar como o contexto político-normativo do setor público influencia o comportamento das organizações e como produz incentivos para que essas tenham um enfoque maior na economia de gastos ou na promoção de mais oportunidades de capacitação, como sugerido pela interpretação dos índices de eficiência observados.

A quantidade de participações não captura aspectos relativos à satisfação dos servidores e percepção de utilidade do treinamento, possibilitado por avaliações de reação, tampouco os resultados de aprendizagem e impacto em nível individual e organizacional. Para uma adequada compreensão do mérito e valor das estratégias adotadas pelas organizações da amostra, medidas dessa natureza seriam fundamentais, uma vez que refletem os objetivos prioritários do programa de treinamento, relativos à sua eficácia e efetividade (Mark et al., 2000; Scriven, 1991; Suchman, 1967). Nesse sentido, destaca-se que as análises realizadas no presente estudo tratam de aspectos relativos às alternativas de provisão de treinamentos, de maneira que se adota a perspectiva da eficiência proposta por Suchman (1967), que atua de forma complementar às avaliações de efetividade, com o objetivo de otimizar os resultados das unidades de treinamento. Logo, assume-se aqui a eficácia e efetividade das ações investigadas, cabendo ressaltar que é de fundamental relevância a averiguação destas perspectivas para que os objetivos do TD&E sejam cumpridos.

Tendo sido apresentada a visão geral e a contextualização da avaliação de eficiência proposta, as próximas seções se dedicarão à descrição dos resultados da Análise Envoltória de Dados de forma mais detalhada, comparando-se as eficiências das organizações estudadas. A análise das organizações ao longo dos anos de 2015 a 2019 nos dois estágios definidos como “Produção” e “Entrega” se dedica à determinar se os recursos disponíveis estão sendo alocados de maneira adequada e se a escala produtiva é apropriada, remetendo à adequação dos retornos do investimento no treinamento – em termos de mais participações – e ao porte das unidades – sua equipe – para execução das atividades de treinamento que lhes são demandada.

4.2 Avaliação comparativa da eficiência das unidades de treinamento

4.2.1 Estágio 1: Produção

Conforme mencionado, o primeiro estágio do modelo de análise consiste na consideração de quantas horas de treinamento foram fornecidas por unidades de treinamento a partir dos seus recursos humanos e financeiros. A amostra conta com 06 organizações, chamadas de DMUs e as suas produtividades e eficiências podem ser melhor apreendidas pelo exame da fronteira de eficiência representada na Figura 31. Tendo em vista que há uma grande diferença de escala entre as organizações – por exemplo as organizações “E” e “D” – e que não houve diferença significativa entre os escores de eficiência entre os modelos CCR e BCC, optou-se pela representação de ambas as fronteiras, demonstrando, ao mesmo tempo, e de forma complementar, a fronteira de retornos constantes, representada pela linha, e a fronteira de retornos variáveis, convexa.

De forma geral, verifica-se que a fronteira de eficiência é definida pelas DMU C6 (i.e. a organização “C”, em 2016) e D8, E5, E6 e E7, consideradas, portanto, eficientes, sendo que C6 se destaca por possuir eficiência sob diferentes pressupostos, uma vez que se situa na convergência das fronteiras de eficiência. As demais DMUs são consideradas ineficientes quanto mais distantes estejam das fronteiras. Dessa forma, verifica-se que o centro do gráfico concentra as DMUs mais ineficientes, com destaque para a organização “F” e outras organizações como “C” e “A” em geral estão próximas da eficiência.

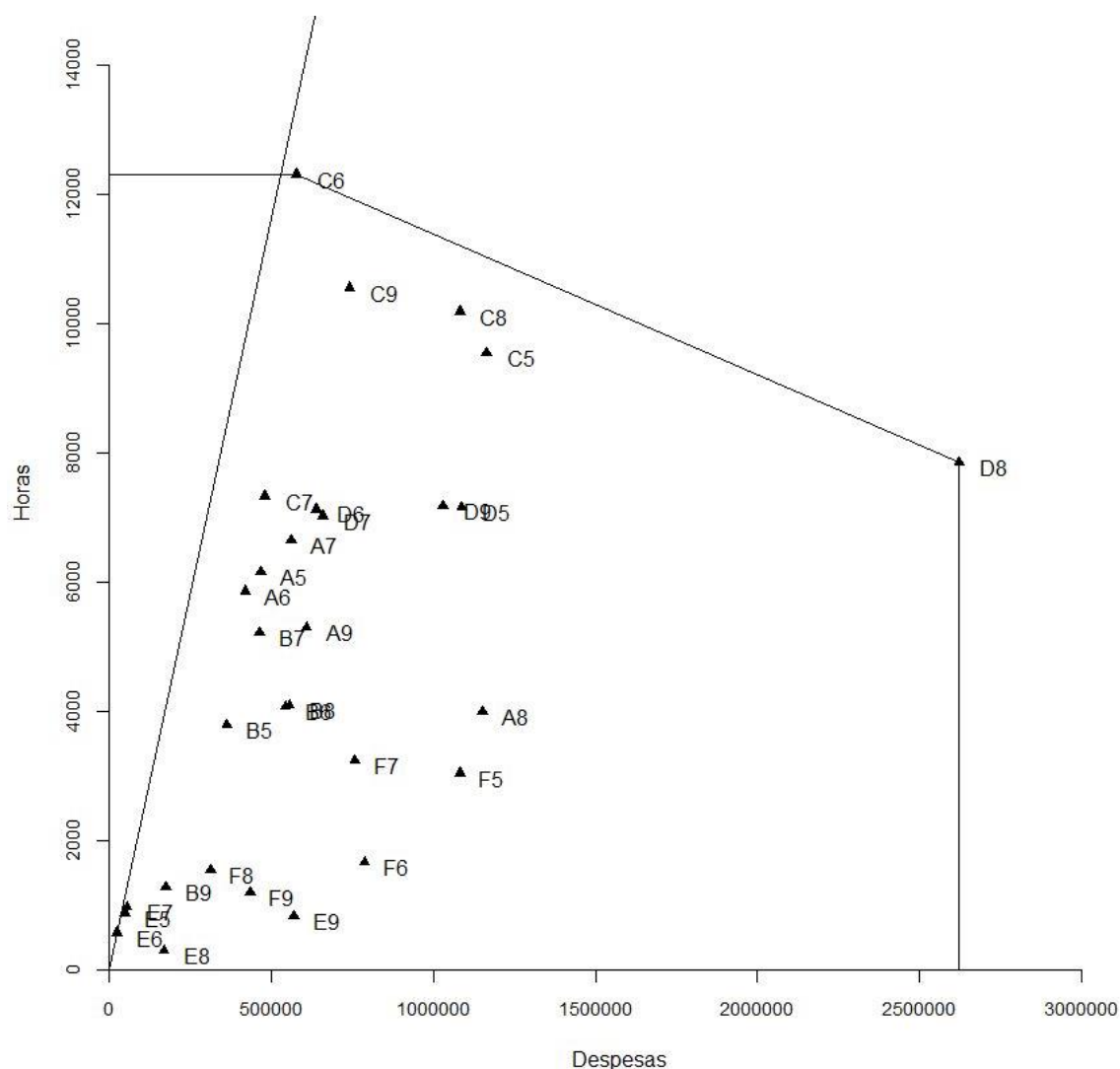


Figura 31. Fronteira de Eficiência CCR x BCC para o Estágio 1

Fonte: Elaboração própria.

A eficiência técnica representa a capacidade da unidade de utilizar uma proporção adequada de insumos para geração de produtos e resulta de sua capacidade gerencial, da qualidade do processo produtivo, da qualificação da força de trabalho e das técnicas e tecnologias adotadas (Ferreira & Gomes, 2020; Mariano, 2007; Rosano-Peña, 2008). Por outro lado, parte da eficiência técnica se deve também à escala adequada de produção (Lovell, 1993). Dessa forma, os escores de eficiência de escala indicam quais organizações operam em um nível ótimo que permita uma conversão eficiente de certa quantidade de insumos em produtos, ou seja, em volume de treinamentos. A Tabela 15 detalha os índices de eficiência técnica e de escala para as DMUs da amostra ao longo do período analisado. Na tabela, são sombreados os valores correspondentes às unidades eficientes e sublinhados os valores com menor eficiência no respectivo período.

Tabela 15
Evolução da Eficiência Técnica e de Escala para o Estágio 1

DMU	2015	2016	2017	2018	2019
<i>Eficiência Técnica – Estágio 01</i>					
A	0,93	0,85	0,84	0,38	0,51
B	0,87	0,78	1,00	0,65	0,38
C	0,78	1,00	0,72	0,93	1,00
D	0,75	0,85	0,83	1,00	0,96
E	1,00	1,00	1,00	0,30	0,86
F	<u>0,41</u>	<u>0,22</u>	<u>0,44</u>	<u>0,26</u>	<u>0,17</u>
<i>Eficiência de Escala – Estágio 01</i>					
A	0,96	<u>0,94</u>	0,98	0,90	1,00
B	0,90	1,00	1,00	0,98	0,90
C	<u>0,79</u>	1,00	1,00	0,86	1,00
D	0,92	1,00	0,99	1,00	0,95
E	0,83	1,00	<u>0,80</u>	<u>0,31</u>	<u>0,24</u>
F	0,95	0,95	0,94	0,89	0,95

Nota. Fonte: Elaboração própria.

Os escores demonstram que, comparativamente, há 07 DMUs que podem ser consideradas tecnicamente eficientes na produção de treinamentos em diferentes períodos. Dessas 07, cinco delas – “C6”, “E6”, “B7”, “D8” e “C9” – operam em uma escala adequada, possuindo tanto eficiência técnica pura, quanto de escala (Lovell, 1993. Rosano-Peña, 2008).

As DMU “E5” “E7” são tecnicamente eficientes, mas apresentam ineficiência de escala, o que indica que o seu porte pode ser inadequado para o seu nível de provisão de treinamentos. A ineficiência de escala se agrava em 2018 e 2019, anos em que a organização possui uma queda acentuada em sua eficiência técnica global.

Quando se analisam os dados de insumo, verifica-se que houve um grande aumento nos recursos financeiros (X2) em 2018 e 2019, após um episódio de acentuada mudança organizacional relatado pelo gestor da unidade. No entanto, não houve aumento da equipe (X1), a qual, por sua vez, é a menor de toda a amostra – 03 servidores. Tal desequilíbrio entre esses fatores pode ter ensejado a dificuldade da equipe em operacionalizar o orçamento ampliado e novas demandas organizacionais para uma oferta adequada de cursos. Nesse sentido, os índices de eficiência técnica e de escala servem de indicativo de necessidades de ajustes técnicos e estruturais nas unidades de treinamento para que os seus recursos possam ser utilizados de forma eficiente, evitando que, conforme apontado por Silva et al. (2016), os recursos providos para unidades ineficientes possam gerar ainda mais desperdícios.

As organizações “A” e “B” também apresentaram uma mudança acentuada de desempenho tendo mantido índices altos entre 2015 e 2017 e sofrendo uma redução acentuada a partir de 2018. As organizações “C” e “D”, por sua vez, possuem um nível de eficiência consistentemente elevado tendo alcançado o escore máximo ou muito elevados em 2016, 2018 e 2019, indicando uma tendência contrária às demais organizações de aumento da eficiência ao longo do tempo. Por fim, nota-se que a organização “F” é consistentemente a unidade mais ineficiente da amostra com relação à aquisição e produção de treinamentos.

Tendo em vista o tamanho da amostra, buscou-se a realização de um procedimento de recálculo dos índices de eficiência por reamostragem aleatória (*bootstrapping*), visando a construção de intervalos de confiança e determinação de quais diferenças entre os índices de eficiência são significativas com índice de confiabilidade de 95%. Um primeiro resultado que a análise DEA com *bootstrap* possibilita é a estimativa de uma nova fronteira de eficiência (Figura 32), com uma fronteira corrigida – linha tracejada em azul – e um intervalo de confiança – linha pontilhada em vermelho (Bogetoft & Otto, 2019).

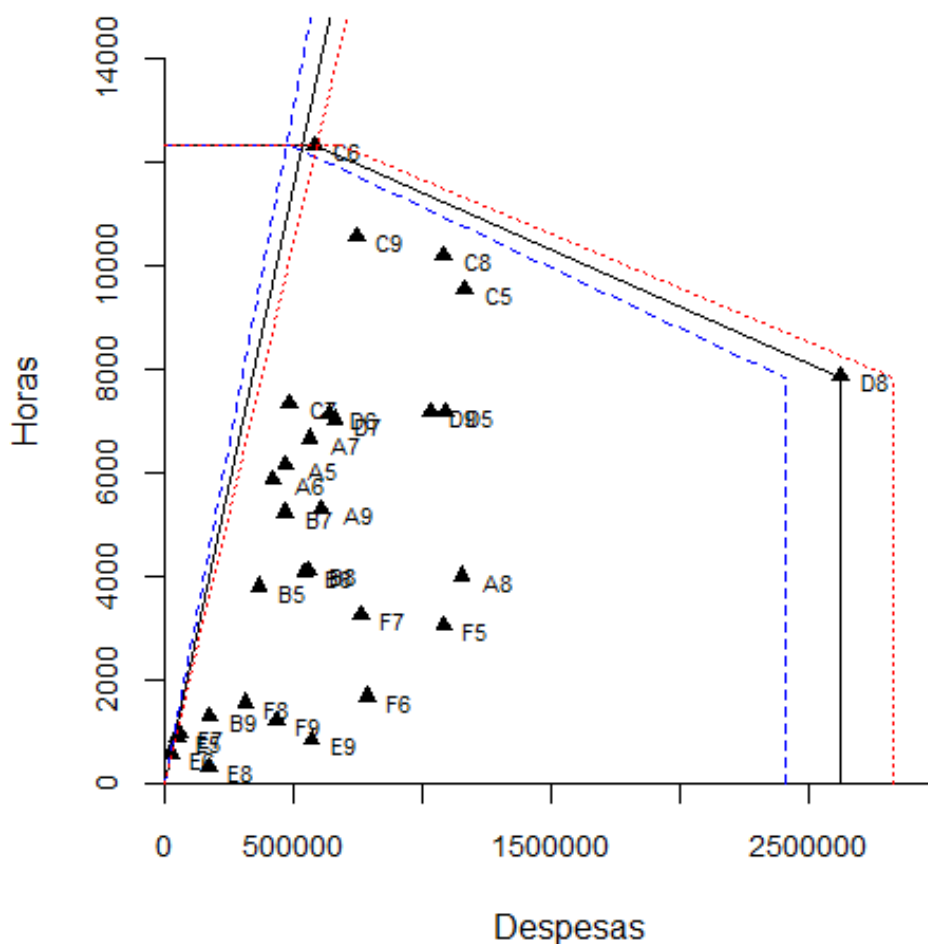


Figura 32. Fronteira de Eficiência CCR x BCC corrigida para o Estágio 1
Fonte: Elaboração própria.

A correção da fronteira enseja a reconsideração dos índices de ineficiência, uma vez que algumas organizações estarão agora mais próximas da nova fronteira, sendo, portanto, possivelmente não tão ineficientes como apontado anteriormente. O exame de cada DMU individualmente, no entanto, é que permite, de fato, a análise mais apurada e a confirmação acerca dos índices de eficiência. A Figura 33 apresenta, dessa forma, as estimativas dos índices de eficiência com intervalos de confiança calculados pelo método *bootstrap* (Bogetoft & Otto, 2019).

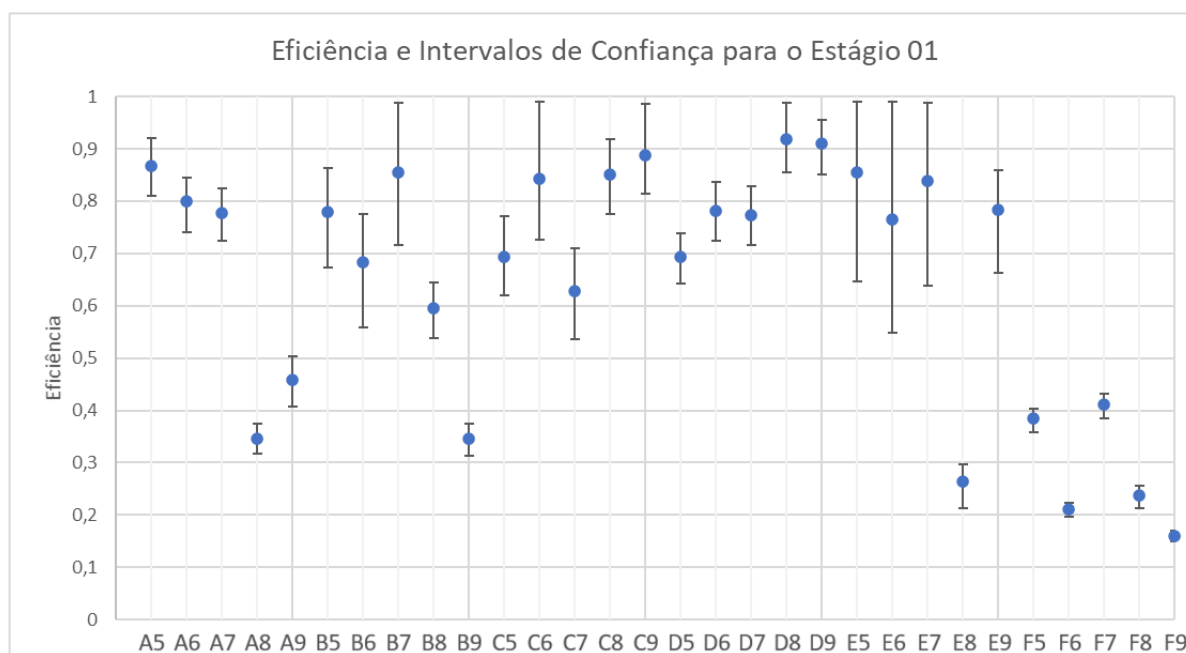


Figura 33. Eficiência e Intervalos de confiança para o Estágio 1

Fonte: Elaboração própria.

O gráfico acima apresenta cada uma das DMUs e a variação para seus índices de eficiência, considerando possíveis erros de estimativa, permitindo uma comparação mais precisa entre as DMUs. Nesse ínterim, novamente a organização “F” pode ser considerada, efetivamente, a mais ineficiente em relação ao restante da amostra. Além disso, as DMUs “A8”, “A9”, “B9” e “E8” também são significativamente mais ineficientes que as demais. Com relação à organização “F”, sua ineficiência decorre antes de questões técnicas e decisões acerca da oferta de treinamentos do que de questões de escala.

Salvo variados fatores que concorrem para a explicação do desempenho dessa unidade, pode-se verificar em seus dados que, por exemplo, nos anos de 2015 (F5) e 2017 (F7), em que houve uma melhoria de seus índices de eficiência, a unidade se valeu de maior diversidade de modalidades e tipos de treinamento para além dos cursos externos em modalidade presencial. Os cursos presenciais contratados junto a empresas somaram apenas 40% da carga horária da

unidade, contra 80% em média nos demais anos. A excessiva contratação de cursos de curta duração junto a empresas, portanto, foi um dos fatores detrimentais à sua eficiência.

Tais explicações suscitam testes mais apurados, mas também devem ser complementadas por outras análises, como por exemplo, a análise de *benchmarking*. A sessão seguinte visa aprofundar a análise da eficiência com relação aos ajustes técnicos que podem ser realizados pelas unidades tendo como referência parceiros de melhores práticas (*peers*), ou seja, organizações que servem de referência para as demais para que se tornem eficientes (Ferreira & Gomes, 2020).

4.2.1.1 Estágio 01: Análise de melhores práticas das unidades eficientes

Uma das premissas do DEA é a definição de quais organizações são mais eficientes na atividade analisada e que podem servir de referência para que outras organizações também se tornem eficientes. Para tanto, foi realizada uma análise de *benchmarks* em que se atribui um valor λ_k aos pares de DMUs eficientes e ineficientes. Quanto maior esse índice, mais importante é a DMU eficiente para determinada DMU ineficiente, variando a partir de 0 e podendo chegar a “1,0”, o que indica que a DMU ineficiente deveria copiar por completo as práticas da DMU de referência. Essa abordagem foi criada com o intuito de parear DMUs com referências factíveis, cuja escala e tecnologia de produção fosse similar. Além do ajuste com base na distância em relação à fronteira de eficiência, as DMUs também podem ajustar imperfeições na própria alocação de seus insumos. Assim, a análise de folgas (*slacks*) visa indicar em que aspectos os recursos estão sendo mal aproveitados, ou seja, sendo empregados sem gerar benefícios adicionais compatíveis com o que se observa na amostra.

Tais informações estão apresentadas na Tabela 16. Nesta, para cada caso apresenta-se o escore de eficiência da DMU e seus insumos e produtos para que se tenha dimensão do ajuste proposto com base na análise de folgas. A última coluna, por sua vez, traz a informação de qual DMU serve de referência de melhor prática para essa unidade, ou seja, aquela cujo pareamento apresentou maior índice λ_k .

Segundo Ferreira & Gomes (2020), quanto mais vezes uma DMU serve de referência para as demais, mais importante na amostra ela é, pois representa a melhor tecnologia para a realização da atividade. Nesse sentido, a DMU “B7” é a mais relevante, pois é referenciada 18 vezes, seguida pelas DMUs “C9” (13), “C6” (12), “D8” (9), “E7” (9), “E6” (1). Ressalte-se que apenas as unidades eficientes são consideradas como referência e que na tabela estão apresentadas apenas as referências com maior índice λ_k . Além disso, tendo em vista que no presente estudo determinadas DMUs são a mesma organização em períodos diferentes, é comum que a DMU de referência seja uma versão eficiente dela mesma em um período distinto.

Esse é o caso das organizações “C” e “B”, mas não é o caso, por exemplo, da organização “D” que, apesar de ter apresentado eficiência em 2018, para os demais anos possui como parceiros de referência as organizações “C” e “B”, denotando que 2018 pode ter sido um período atípico para “D”, caso contrário serviria como referência para os demais anos.

Tabela 16
Análise de *Benchmarks* para o Estágio 1

Caso	Eficiência	X1	X2	Z	Ajuste X1	Ajuste X2	Ajuste Z	Referência
A5	0,93	6	468.519,34	6.143	0		0	B7 (0,72)
A6	0,85	7	420.189,08	5.852	0		0	B7 (0,44)
A7	0,83	7	560.260,28	6.642	0		0	B7 (0,55)
A8	0,38	9	1.151.019,56	3.990	0	409.177,83	0	C9 (1,0)
A9	0,51	10	608.249,91	5.285	0		0	C6 (0,49)
B5	0,87	4	363.309,08	3.780	0		0	B7 (0,72)
B6	0,78	4	543.155,74	4.065	0	79.713,77	0	B7 (1,0)
B7	1,00	4	463.441,97	5.214	0		0	-
B8	0,65	5	555.547,87	4.090	0	-	0	B7 (0,78)
B9	0,38	5	173.533,45	1.273	0	-	0	E7 (0,76)
C5	0,78	13	1.163.696,07	9.540	0	586.255,06	0	C6 (1,0)
C6	1,00	13	577.441,01	12.305	0	-	0	-
C7	0,72	12	479.808,00	7.324	0,7647	-	0	C6 (0,82)
C8	0,93	10	1.081.035,36	10.177	0	380.293,81	0	C9 (0,75)
C9	1,00	9	741.841,73	10.546	0	-	0	-
D5	0,75	8	1.087.947,59	7.153	0	-	0	C9 (0,72)
D6	0,85	7	639.785,16	7.121	0	-	0	C9 (0,60)
D7	0,83	7	658.862,97	7.018	0	-	0	C9 (0,60)
D8	1,00	6	2.622.944,43	7.843	0	-	0	-
D9	0,96	6	1.028.733,82	7.169	0	-	0	B7(0,46)
E5	1,00	3	47.820,00	873	0	-	0	-
E6	1,00	3	24.190,00	562	0	-	0	-
E7	1,00	3	54.453,60	958	0	-	0	-
E8	0,30	3	168.427,00	286	0	113.973,40	0	E7 (1,0)
E9	0,86	3	568.303,10	826	0	513.849,50	0	E7 (1,0)
F5	0,41	6	1.083.284,15	3.036	0	-	0	B7 (0,45)
F6	0,22	6	787.210,30	1.657	0	-	0	B7 (0,54)
F7	0,44	6	758.105,98	3.224	0	-	0	B7 (0,55)
F8	0,26	7	312.577,69	1.533	0	-	0	E7 (0,47)
F9	0,17		433.888,83	1.192	0	-	0	B7 (0,48)

Nota. Fonte: Elaboração própria.

Já em relação à análise de folgas, 07 organizações apresentam recursos que podem ser ajustados sem prejuízo de suas atividades, nenhuma delas eficiente, no entanto. A organização “E” em 2018 (E8) e (E9), por exemplo, dispendeu recursos financeiros em excesso para a obtenção de um resultado comparativamente menor, tendo em vista seu desempenho anterior em 2017 (E7), de forma que ela deveria observar como foi o seu programa de treinamento em 2017 para buscar replicar as estratégias bem-sucedidas.

Porém, deve-se ter em mente que isso se refere a um critério estritamente relacionado ao quantitativo de treinamento promovido, o que não remete à sua qualidade. É possível que os recursos tenham gerado outros tipos de retornos não considerados na análise, de natureza qualitativa e mais efetivos como, por exemplo, cursos mais avançados e estratégicos, ainda que mais caros. A DMU “A8”, por exemplo, poderia ter utilizado apenas 64% dos recursos financeiros gastos para obter o mesmo resultado em horas de capacitação, porém é possível que mesmo com menos horas de capacitação, os cursos, de maior valor, tenham visado atender outros critérios como os acima sugeridos.

As DMUs “B6”, “C5” e “C8” apresentam casos que devem ser analisados com cautela, pois seus ajustes têm por referência suas versões futuras “B7”, “C6” e “C9”, respectivamente. Ou seja, as organizações, na prática, melhoraram seu desempenho de um ano para o outro. Já em “C7”, verifica-se um ajuste de redução de equipe, tendo em vista o seu desempenho em “C6”. Logo, é relevante, em uma análise intertemporal, observar se os *benchmarks* se referem a situações passadas ou futuras, para que se limite ou adeque a interpretação acerca do que são as melhores práticas. Pode ser que determinada organização, à sua época, fosse a melhor prática e tenha sido considerada ineficiente em razão de avanços que apenas DMUs futuras tiveram acesso.

Por fim, de forma geral, as folgas e os *benchmarks* não estão necessariamente interligados. As folgas devem ser interpretadas, no presente contexto de pesquisa, como indicativos de situações de gastos desvantajosos para determinada quantidade de horas de treinamentos, enquanto os *benchmarks* servem ao propósito de fornecer um “alvo” para a melhoria da eficiência, sendo mais significativos para a tomada de decisão do que a simples consideração de ajustes de recursos financeiros (Ferreira & Gomes, 2020).

Apesar dos cuidados trazidos pelos modelos de estimação de intervalos e a consideração de múltiplas referências de melhores práticas, caso se deseje realizar um ranqueamento dentre as próprias organizações eficientes, uma alternativa é a realização de uma análise de supereficiência (Ferreira & Gomes, 2020). A Tabela 17 traz uma breve análise que permite ordenar as 07 DMUs eficientes nesse estágio do processo de treinamento. A análise permite

estimar um valor acima do limite de “1,0” para eficiência. O valor “Inf” corresponde a uma tendência ao infinito, mas por não precisar um escore, será considerado como um valor desconhecido. Sendo assim, a DMU “C6” pode ser considerada como o melhor desempenho, em termos de eficiência de aquisição e produção de treinamentos, dentre as demais da amostra, com escore estimado de 1,44.

Tabela 17
Análise de Supereficiência para o Estágio 1

DMU	Supereficiência
C6	1,44
B7	1,32
C9	1,13
E7	1,10
D8	1,07
E5	1,00
E6	Inf

Nota. Fonte: Elaboração própria.

Logo, em toda a amostra, o programa de treinamento executado pela organização “E” no ano de 2016 foi o melhor em termos de utilização de recursos para a oferta de maior quantidade de treinamentos. Com feito, o ano de 2016 foi o que a organização “C” ofereceu o maior número de cursos e horas de treinamento em todo o período analisado. Conforme relatado pelos gestores, 2016 corresponde ao período de suspensão do orçamento de capacitação em todo o setor público federal. Dessa forma, o indicador pode refletir uma situação atípica de excessiva contenção de despesas, ao mesmo tempo que demonstra o sucesso da organização “C” em aumentar o fluxo de treinamentos, mesmo diante das condições adversas de seu contexto.

4.2.2 Estágio 2: Entrega

O Estágio 2 visa capturar a capacidade das unidades de ofertar ações de treinamento para o maior número de servidores, com o quantitativo de horas de treinamento implementadas. Sendo assim, devido à orientação ao produto (*output*) utilizada na análise, são privilegiadas estratégias de fornecimento de turmas com maior quantidade de alunos. Os tipos de ações educacionais com essas características podem ser, por exemplo, eventos como palestras e seminários de curta duração oferecidos internamente pelas organizações ou os chamados cursos fechados, em que são contratadas ou planejadas e implementadas turmas exclusivas para grupos de servidores da organização.

Dessa forma, o indicador visa identificar organizações que planejam a maior parte de suas ações organizacionais, em contraste com aquelas organizações que privilegiam contratações de vagas avulsas em cursos de mercado. Além disso, se consideradas as prescrições da literatura de TD&E acerca da etapa de “entrega” do treinamento, a maior eficiência nesse indicador também pode referir-se à capacidade da unidade de treinamento de comunicar adequadamente as oportunidades de treinamento e de conhecer melhor as características de seu público-alvo promovendo maior adesão aos cursos ofertados (Borges-Andrade et al., 2012; Salas et al., 2012). Assim, tendo como objetivo a provisão equânime de treinamentos a um maior número de servidores, considera-se que a abordagem adotada na análise atende ao objetivo de apoiar a gestão dessa atividade de forma alinhada aos pressupostos do campo de TD&E.

Para o Estágio 2 os pressupostos de retornos variáveis se apresentaram como mais adequados para os dados da amostra. Nesse sentido, a representação da fronteira de eficiência (Figura 34) será distinta da apresentada para o Estágio 1. Como já observado anteriormente, as DMUs são menos eficientes no Estágio 02 do que no Estágio 01. Dessa forma, verifica-se que a fronteira de eficiência é definida apenas pelas DMUs “F6”, “C7” e “E8”. Verifica-se também que o centro do gráfico abriga os conjuntos de DMUs ineficientes, sendo, porém, as mais ineficientes as relativas à organização “B”.

Um aspecto relevante é o fato da organização “F”, de pior desempenho no Estágio 01, possuir índices consistentes de alta eficiência no Estágio 02. Da mesma forma, “E8” considerada eficiente nesse estágio apresenta um dos piores desempenhos no Estágio 01, o que pode ser uma evidência de que essas organizações podem ter realizado uma estratégia própria em que se buscou o fornecimento de treinamentos de maior custo, mas com maior alcance.

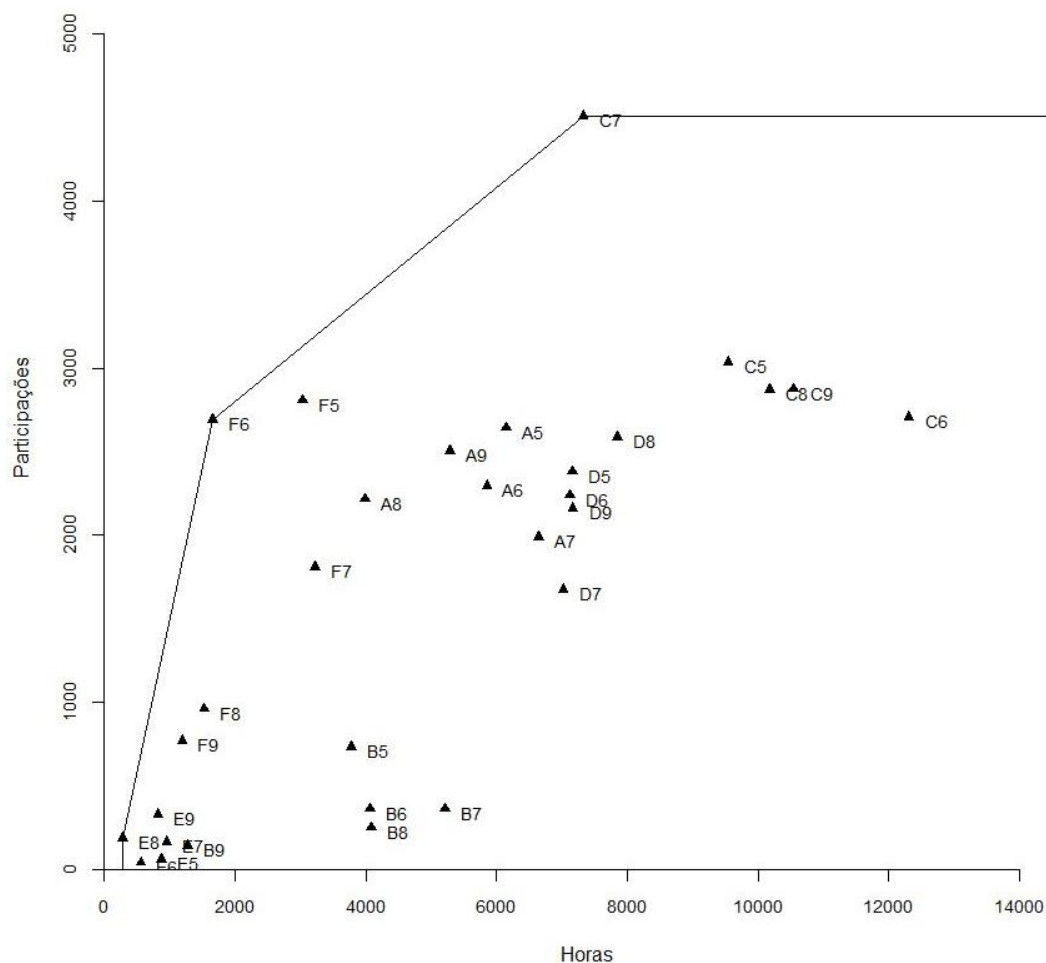


Figura 34. Fronteira de Eficiência BCC sem orientação específica para o Estágio 2
Fonte: Elaboração própria.

Quando verificados os índices de eficiência técnica na Tabela 18, nota-se que as organizações “C” e “D” mantêm um desempenho mediano, mas ainda superior às unidades “B” e “E”, com índices muito baixos na maior parte dos períodos. Boa parte da ineficiência das DMUs, no entanto, pode ser atribuída a problemas de escala e não a questões técnicas do processo de treinamento.

Tabela 18

Evolução da Eficiência Técnica e de Escala para o Estágio 2

DMU	2015	2016	2017	2018	2019
<i>Eficiência Técnica – Estágio 02</i>					
A	0,64	0,57	0,46	0,64	0,65
B	0,22	0,10	0,09	0,07	0,07
C	0,67	0,60	1,00	0,64	0,64
D	0,53	0,50	0,38	0,57	0,48
E	0,05	0,05	0,11	1,00	0,28
F	0,90	1,00	0,57	0,39	0,42

Eficiência de Escala – Estágio 02					
A	0,41	0,42	0,40	0,53	0,45
B	0,55	0,52	0,45	0,52	0,96
C	0,29	0,23	0,38	0,27	0,26
D	0,38	0,38	0,39	0,35	0,38
E	0,89	0,76	0,91	0,40	0,87
F	0,64	1,00	0,61	0,99	0,95

Nota. Fonte: Elaboração própria.

Dessa forma, as unidades parecem ser bem adequadas para atividades de aquisição e produção de treinamentos (Estágio 01), mas talvez não possuam uma escala ótima quando se trata de promover o maior alcance da capacitação em toda a organização (Estágio 02). Cabe refletir, no entanto, que ajustes seriam necessários com relação à escala, sendo possível, conforme a literatura, que se trate de uma decisão entre a criação de unidades de treinamento menores e com maior capilaridade na organização ou grandes centros de treinamento centralizados. Para uma distinção adequada das eficiências, procedeu-se também para o Estágio 02 à utilização de técnicas de *bootstrapping*, para construção de intervalos de confiança. A estimativa da fronteira de eficiência corrigida pode ser verificada na Figura 35.

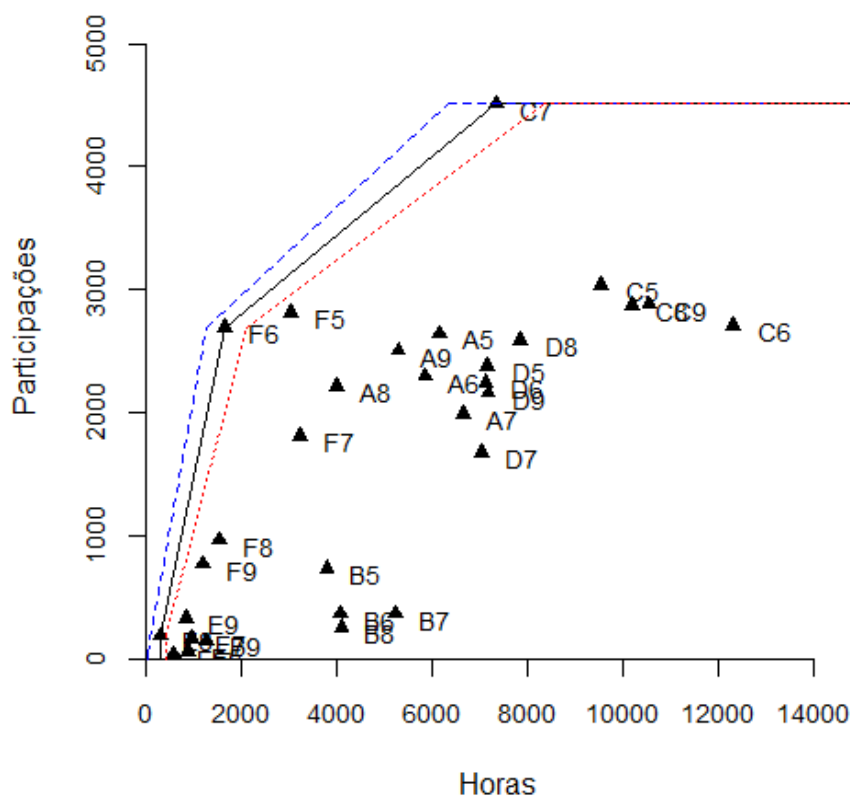


Figura 35. Fronteira de Eficiência BCC corrigida para o Estágio 2
Fonte: Elaboração própria.

Com relação à análise de eficiência das DMUs, a Figura 36 demonstra que a organização “B” apresenta, efetivamente, baixos índices de eficiência nessa etapa do processo de treinamento. A organização “F”, apesar de um bom desempenho médio, apresenta um declínio ao longo do período de análise. Dentre as explicações possíveis, a análise dos dados da unidade indica que em “F8” e “F9” um montante muito menor de investimento em treinamento foi realizado e que esse se deu quase que completamente por meio de ações presenciais contratadas junto a empresas, reforçando o já mencionado de que esse indicador tende a ser menor para as organizações que adotam essa estratégia. Com relação à organização “E”, como já abordado anteriormente, observa-se, aqui, também, uma mudança drástica a partir de 2018, em que a unidade recebeu um significativo incremento nos investimentos de treinamento, o que pode ter se refletido na melhoria de seus resultados de aumento de participações em treinamento.

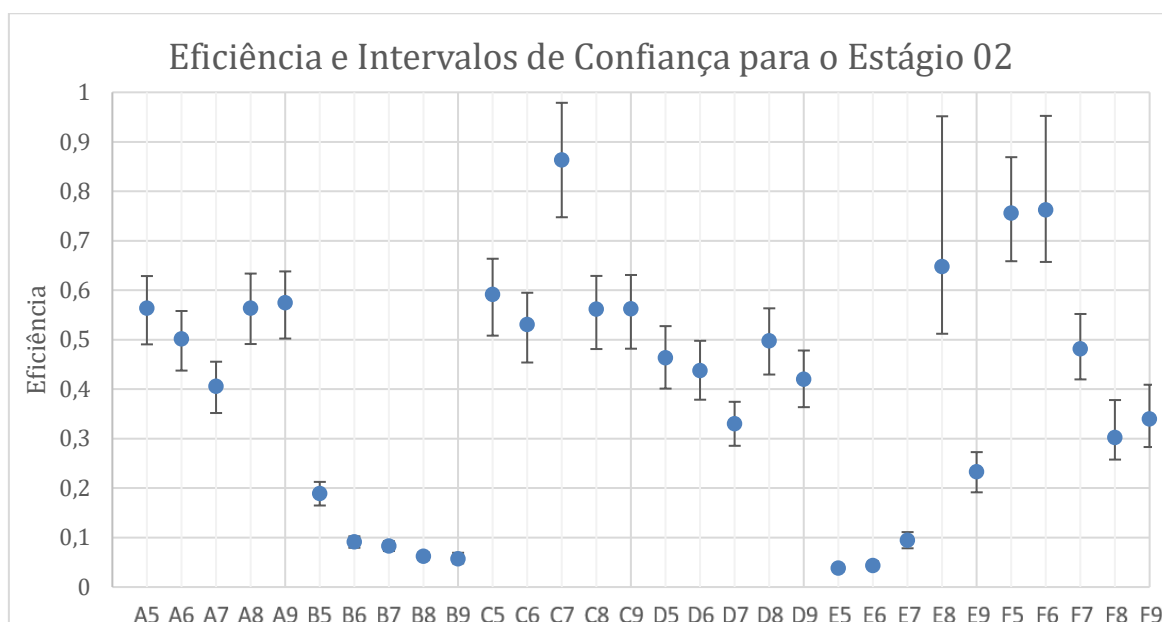


Figura 36. Eficiência e Intervalos de confiança para o Estágio 2

Fonte: Elaboração própria.

Analogamente ao realizado para o Estágio 1, a sessão seguinte visa aprofundar a análise da eficiência com relação aos ajustes técnicos que podem ser realizados pelas unidades tendo como referência parceiros de melhores práticas (*peers*), ou seja, organizações que servem de referência para as demais para que se tornem eficientes (Ferreira & Gomes, 2020).

4.2.2.1 Estágio 02: Análise de melhores práticas das unidades eficientes

De forma similar ao Estágio 01, a análise de folgas e *benchmarks* visa orientar os ajustes a serem realizados pelas DMUs ineficientes, os quais tendem a ser mais significativos nesse Estágio 02 com poucos casos eficientes. A Tabela 19, portanto, apresenta os ajustes nas horas

de treinamento (Z). Primeiramente, nota-se que as folgas se concentram na organização “C”, a qual possui um volume de oferta de cursos muito superior às demais. Nesse sentido, a análise de folgas pode adotar diversas interpretações para esses resultados. Uma possibilidade é que existe um limite a partir do qual os indivíduos não possam participar dos treinamentos, ou seja, as folgas são um excedente de uma oferta ótima para o quantitativo de servidores. Outra perspectiva é de que estão sendo ofertadas ações educacionais que não promovem um retorno adequado de participantes, o qual pode se dever à abrangência limitada do treinamento perante o público-alvo, ou mesmo ao baixo comparecimento e falta de interesse dos indivíduos em oportunidades de treinamento ofertadas.

Tabela 19
Análise de *Benchmarks* para o Estágio 2

Caso	Eff.	(Z)	(Y)	Folgas Z	Folgas Y	Referência
A5	0,64	6.143	2643	0	R\$ -	C7 (0,79)
A6	0,57	5.852	2294	0	R\$ -	C7 (0,74)
A7	0,46	6.642	1988	0	R\$ -	C7 (0,88)
A8	0,64	3.990	2214	0	R\$ -	F6 (0,59)
A9	0,65	5.285	2504	0	R\$ -	C7 (0,64)
B5	0,22	3.780	732	0	R\$ -	F6 (0,63)
B6	0,10	4.065	360	0	R\$ -	F6 (0,58)
B7	0,09	5.214	361	0	R\$ -	C7 (0,63)
B8	0,07	4.090	247	0	R\$ -	F6 (0,57)
B9	0,07	1.273	141	0	R\$ -	F6 (0,72)
C5	0,67	9.540	3036	2216	R\$ -	C7 (1,0)
C6	0,60	12.305	2707	4981	R\$ -	C7 (1,0)
C7	1,00	7.324	4510	0	R\$ -	-
C8	0,64	10.177	2872	2853	R\$ -	C7 (1,0)
C9	0,64	10.546	2875	3222	R\$ -	C7 (1,0)
D5	0,53	7.153	2379	0	R\$ -	C7 (0,97)
D6	0,50	7.121	2239	0	R\$ -	C7 (0,96)
D7	0,38	7.018	1673	0	R\$ -	C7 (0,95)
D8	0,57	7.843	2585	519	R\$ -	C7 (1,0)
D9	0,48	7.169	2160	0	R\$ -	C7 (0,97)
E5	0,05	873	58	0	R\$ -	E8 (0,57)
E6	0,05	562	37	0	R\$ -	E8 (0,80)
E7	0,11	958	161	0	R\$ -	E8 (0,50)
E8	1,00	286	186	0	R\$ -	-
E9	0,28	826	326	0	R\$ -	E8 (0,60)

F5	0,90	3.036	2808	0	R\$ -	F6 (0,75)
F6	1,00	1.657	2690	0	R\$ -	-
F7	0,57	3.224	1807	0	R\$ -	F6 (0,72)
F8	0,39	1.533	958	0	R\$ -	F6 (0,90)
F9	0,42	1.192	768	0	R\$ -	F6 (0,66)

Nota. Fonte: Elaboração própria.

Com relação às melhores práticas, a DMU “F6” é referenciada 23 vezes, sendo, portanto, o caso de maior importância, seguido pela DMU “C7”, referenciada 21 vezes e “E8”, 08 vezes. Novamente, a organização “C” é referência para si mesma no período de 2017 e é a principal referência para a organização “D”. Como ambas atuam em setores de governo similares, é possível, portanto, que essas organizações tenham sistemas e processos de treinamento similares, o que levanta questões sobre a possibilidade de trocas de conhecimentos e experiências, que é, justamente o objetivo da aplicação prática do DEA.

Para as demais organizações, em geral, essas possuem como parceiro de melhores práticas versões suas, com exceção das organizações “A” e “B” que se dividem entre as DMUs “C7” e “F6” como referência. Dentre essas, verifica-se, com base na análise de supereficiência (Tabela 20), que “F6” pode ser considerada a DMU com melhor desempenho, em termos de eficiência na entrega de treinamentos ao maior número de servidores.

Tabela 20
Análise de Supereficiência para o Estágio 2

DMU	Supereficiência
F6	1,80
C7	1,52
E8	Inf

Nota. Fonte: Elaboração própria.

Pode-se notar que no ano de 2016, apesar das relatadas restrições orçamentárias por parte do governo, ocorreram os dois programas de treinamento mais bem-sucedidos da amostra do ponto de vista dos critérios adotados, ainda que em perspectivas distintas. A organização “C” em 2016 obteve o melhor desempenho no Estágio 1, enquanto a organização “F” em 2016 destacou-se em uma estratégia distinta. Verificando-se os registros de capacitação, verifica-se que “F” teve o maior número de participações em 2016, tendo realizado variados cursos para capacitação de quantitativos entre 130 e 150 servidores e um programa nacional de palestras que alcançou 600 pessoas, o que pode ter ensejado o alto valor do índice. Porém, a mesma estratégia não foi adotada nos anos seguintes de 2017 a 2019, de forma que a organização “F”

demonstra uma deterioração da sua eficiência com uma redução considerável no quantitativo de servidores capacitados de cerca de 2.011 pessoas em 2016 para apenas 595 e 571 em 2018 e 2019. Dessa forma, os índices de eficiência desse Estágio foram capazes de capturar comparativamente essa queda no desempenho da unidade.

4.2.3 Eficiência Global

A eficiência global consiste em um dos principais objetivos metodológicos deste estudo e busca representar a eficiência de unidades de treinamento na provisão de ações de capacitação para o maior número de pessoas com o menor custo em termos de recursos humanos e financeiros. Tendo em vista o modelo DEA em rede utilizado nesta pesquisa, a eficiência global é calculada pela multiplicação dos índices de eficiência dos dois estágios, definidos com base em variáveis sistematicamente levantadas junto às organizações para análise do processo de treinamento: os estágios de “Produção” e “Entrega” (Cook et al., 2010; Ferreira & Gomes, 2020).

A partir da descrição e compreensão de cada um dos estágios constituintes do modelo de análise, realizada nas seções anteriores, torna-se possível a apresentação e interpretação dos resultados de eficiência global. A Tabela 21 apresenta os escores de eficiência global para cada uma das DMUs analisadas, ordenando-as em ordem decrescente a partir da mais eficiente. Juntamente com o escore de eficiência global são resgatados os respectivos índices de eficiência dos Estágios 01 e 02, assim como os valores dos insumos e produtos cujas combinações resultaram nos índices apresentados. Com isso, torna-se possível compreender quais dessas combinações foram consideradas eficientes nos diferentes estágios de análise e como se refletem de forma conjunta no escore de eficiência global.

Por resultarem do produto dos escores de eficiência nos estágios de treinamento, observou-se que nenhuma DMU foi considerada completamente eficiente, sendo o maior escore de eficiência global encontrado de “0,72” para a DMU “C7”. A DMU apresentou eficiência igual a “0,72” na primeira etapa e escore de “1,0” na segunda etapa, o que definiu o valor do escore global. A mesma organização “C” também foi considerada a segunda e terceira mais eficiente com “C9” e “C6”, podendo ser considerada, portanto, a organização com as melhores práticas globais para execução de processos de treinamento dentre as organizações da amostra. As organizações “A” e “D” seguem nas posições seguintes com eficiências globais, em sua maior parte, acima do valor médio da amostra de aproximadamente “0,30”.

Tabela 21

Eficiência global e combinações de insumos e produtos

N	DMU	Eficiência Global	Eficiência Produção	Eficiência Entrega	X1 (Equipe)	X2 (Despesas)	Z (Horas)	Y (Part.)
1	C7	0,72	0,72	1,00	12	479.808,00	7.324,00	4.510,00
2	C9	0,64	1,00	0,64	9	741.841,73	10.546,00	2.875,00
3	C6	0,60	1,00	0,60	13	577.441,01	12.305,00	2.707,00
4	A5	0,59	0,93	0,64	6	468.519,34	6.143,00	2.643,00
5	C8	0,59	0,93	0,64	10	1.081.035,36	10.177,00	2.872,00
6	D8	0,57	1,00	0,57	6	2.622.944,43	7.843,00	2.585,00
7	C5	0,52	0,78	0,67	13	1.163.696,07	9.540,00	3.036,00
8	A6	0,48	0,85	0,57	7	420.189,08	5.852,00	2.294,00
9	D9	0,47	0,96	0,48	6	1.028.733,82	7.169,00	2.160,00
10	D6	0,43	0,85	0,50	7	639.785,16	7.121,00	2.239,00
11	D5	0,40	0,75	0,53	8	1.087.947,59	7.153,00	2.379,00
12	A7	0,39	0,83	0,46	7	560.260,28	6.642,00	1.988,00
13	F5	0,36	0,41	0,90	6	1.083.284,15	3.036,00	2.808,00
14	A9	0,33	0,51	0,65	10	608.249,91	5.285,00	2.504,00
15	D7	0,32	0,83	0,38	7	658.862,97	7.018,00	1.673,00
16	E8	0,30	0,30	1,00	3	168.427,00	286,00	186,00
17	F7	0,25	0,44	0,57	6	758.105,98	3.224,00	1.807,00
18	A8	0,24	0,38	0,64	9	1.151.019,56	3.990,00	2.214,00
19	E9	0,24	0,86	0,28	3	568.303,10	826,00	326,00
20	F6	0,22	0,22	1,00	6	787.210,30	1.657,00	2.690,00
21	B5	0,19	0,87	0,22	4	363.309,08	3.780,00	732,00
22	E7	0,11	1,00	0,11	3	54.453,60	958,00	161,00
23	F8	0,10	0,26	0,39	7	312.577,69	1.533,00	958,00
24	B7	0,09	1,00	0,09	4	463.441,97	5.214,00	361,00
25	B6	0,08	0,78	0,10	4	543.155,74	4.065,00	360,00
26	F9	0,07	0,17	0,42	7	433.888,83	1.192,00	768,00
27	E6	0,05	1,00	0,05	3	24.190,00	562,00	37,00
28	B8	0,05	0,65	0,07	5	555.547,87	4.090,00	247,00
29	E5	0,05	1,00	0,05	3	47.820,00	873,00	58,00
30	B9	0,03	0,38	0,07	5	173.533,45	1.273,00	141,00

Nota. Fonte: Elaboração própria.

As três organizações “B”, “E” e “F” apresentaram os escores globais mais baixos. Mesmo tendo sido consideradas como *benchmarks* de eficiência em uma das duas etapas – e.g. “E6”, “E7”, “B7”, “F6” – o baixo desempenho na outra etapa levou à colocação dessas organizações nas últimas posições do ordenamento. Tal aspecto demonstra, por um lado, o caráter determinístico do DEA e como ele pode se ampliar no modelo em rede, no qual uma

etapa pode determinar um escore global muito baixo, mesmo com um desempenho excelente nas outras etapas. Por outro lado, esse aspecto pode ser considerado também uma vantagem por ser capaz de evidenciar com clareza onde se encontram os gargalos e quais são as principais lacunas e problemas operacionais de unidades produtivas, projetando como eficientes apenas aquelas que apresentem um desempenho efetivo como um todo.

Para ilustrar essa questão, a Tabela 22 associa os escores de eficiência a indicadores de resultado que representam, por exemplo, objetivos de economicidade – custo médio da hora-aula na DMU (Custo H/A) – e de alcance das ações educacionais – proporção da força de trabalho capacitada (%Cap.). Tais critérios são comuns em avaliações de eficiência de programas em termos de objetivos intermediários (Suchman, 1967), sendo objetivos de avaliação também operacionalizados na literatura da Educação (De Witte & López-Torres, 2017) e de eficiência do gasto público (Siqueira et al., 2020). Adicionalmente, um aspecto que se buscará discutir em vista dos índices de eficiência são os tipos e modalidades de treinamento mais utilizados que serão discutidos adiante.

Tabela 22

Eficiência global e indicadores do processo de treinamento

N	DMU	Eficiência Global	Eficiência Produção	Eficiência Entrega	Custo H/A	%Cap.	% Pres. Externo	% Pres. Interno	% EAD
1	C7	0,72	0,72	1,00	65,51	0,94	64%	10%	26%
2	C9	0,64	1,00	0,64	70,34	0,98	54%	4%	42%
3	C6	0,60	1,00	0,60	46,93	0,84	68%	7%	26%
4	A5	0,59	0,93	0,64	76,27	0,49	64%	4%	31%
5	C8	0,59	0,93	0,64	106,22	0,93	62%	3%	35%
6	D8	0,57	1,00	0,57	334,43	0,83	59%	16%	26%
7	C5	0,52	0,78	0,67	121,98	0,87	79%	4%	16%
8	A6	0,48	0,85	0,57	71,80	0,36	50%	7%	43%
9	D9	0,47	0,96	0,48	143,50	0,73	83%	5%	12%
10	D6	0,43	0,85	0,50	89,84	0,82	77%	4%	19%
11	D5	0,40	0,75	0,53	152,10	0,84	86%	6%	8%
12	A7	0,39	0,83	0,46	84,35	0,32	60%	4%	36%
13	F5	0,36	0,41	0,90	356,81	0,61	42%	8%	49%
14	A9	0,33	0,51	0,65	115,09	0,67	40%	6%	53%
15	D7	0,32	0,83	0,38	93,88	0,74	73%	6%	22%
16	E8	0,30	0,30	1,00	588,91	0,13	79%	21%	0%
17	F7	0,25	0,44	0,57	235,14	0,45	43%	4%	53%
18	A8	0,24	0,38	0,64	288,48	0,56	62%	9%	29%

19	E9	0,24	0,86	0,28	688,02	0,28	97%	3%	0%
20	F6	0,22	0,22	1,00	475,08	0,65	59%	9%	32%
21	B5	0,19	0,87	0,22	96,11	0,24	81%	13%	6%
22	E7	0,11	1,00	0,11	56,84	0,14	58%	12%	30%
23	F8	0,10	0,26	0,39	203,90	0,21	82%	17%	1%
24	B7	0,09	1,00	0,09	88,88	0,16	91%	4%	5%
25	B6	0,08	0,78	0,10	133,62	0,13	90%	5%	5%
26	F9	0,07	0,17	0,42	364,00	0,21	94%	5%	2%
27	E6	0,05	1,00	0,05	43,04	0,03	73%	0%	27%
28	B8	0,05	0,65	0,07	135,83	0,12	87%	4%	9%
29	E5	0,05	1,00	0,05	54,78	0,04	100%	0%	0%
30	B9	0,03	0,38	0,07	136,32	0,08	91%	6%	2%

Nota. Fonte: Elaboração própria.

Muito embora as DMUs “E5”, “E6”, “E7” e “B7”, por exemplo, tenham sido eficientes no primeiro estágio de aquisição e produção de ações educacionais, os cursos fornecidos abrangeram, respectivamente, 4%, 3%, 14% e 16% do total de servidores em suas organizações, valores muito inferiores à média de 48% de abrangência da amostra. Também constaram das últimas posições organizações com elevados custos médios por hora-aula, cuja média em toda a amostra é de R\$ 183,93. Em contraste, as DMUs com escores de eficiência global mais elevados apresentavam alcance acima da média, em sua maioria, chegando a 94% e 98% dos servidores de toda a organização capacitados. Com relação aos custos, também se observa que, em sua maioria, são DMU com custos médios bastante inferiores à média observada na amostra. Para melhor ilustrar tais aspectos, buscou-se representar por meio de gráficos de dispersão a relação entre os índices de eficiência global e os citados indicadores de abrangência da força de trabalho (Figura 37) e custo médio da hora-aula (Figura 38).

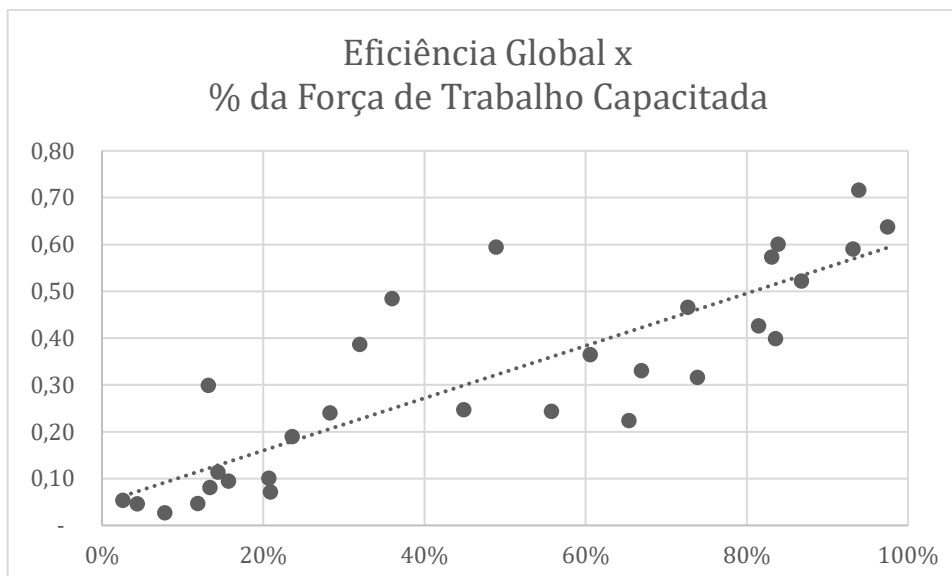


Figura 37. Gráfico Eficiência global x Força de Trabalho capacitada
Fonte: Elaboração própria.

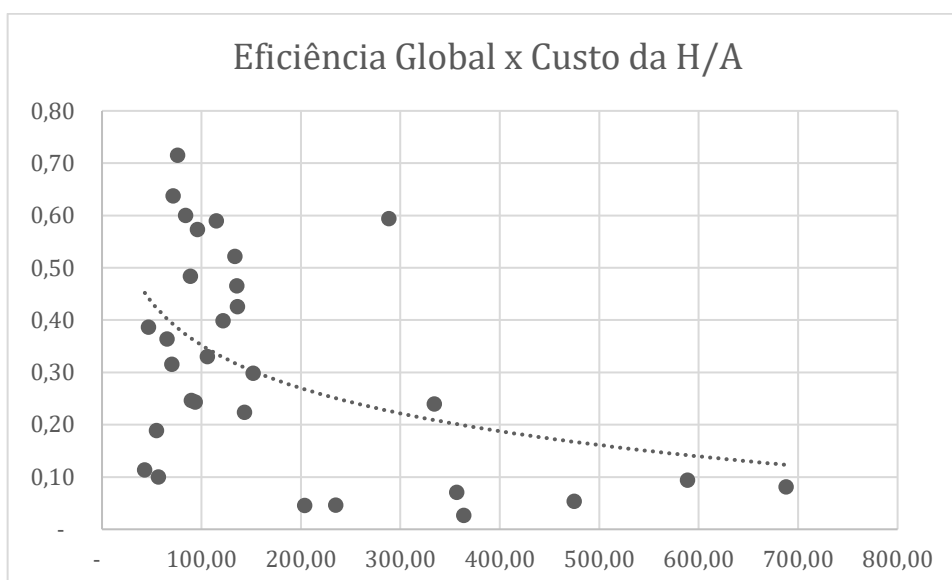


Figura 38. Gráfico Eficiência global x Custo da hora-aula
Fonte: Elaboração própria.

O primeiro gráfico representa, para a amostra analisada, a relação positiva entre os escores de eficiência global e o percentual de servidores capacitados. No segundo gráfico, verifica-se que as organizações com índices de eficiência global elevados se concentram na região correspondente a custos de hora-aula inferiores à média da amostra, estabelecendo uma relação de tendência negativa entre eficiência e custo, que, embora seja esperado, reforça a capacidade do escore global de congregar ambos os resultados.

Dessa forma, o modelo de análise, efetivamente, classifica as organizações por sua capacidade de atender, simultaneamente, a dois critérios de eficiência, quais sejam: fornecer cursos pelo menor custo e para o maior número de servidores, e que termina por congrega, em

um único índice, dois aspectos que correspondem a objetivos de avaliação relevantes no contexto da administração pública federal. A política de capacitação recém-publicada pelo governo federal define como objetivos almejados a capacitação equânime com oferta de oportunidades de treinamento para todos os servidores, com atenção ao melhor custo-benefício (Brasil, 2019a, Camões & Meneses, 2016). Diante disso, o índice de eficiência global pode servir de referência enquanto variável-critério para o desempenho de unidades de treinamento no que se refere à capacidade de implementação da Política Nacional de Desenvolvimento de Pessoas segundo seus objetivos.

A partir do índice de eficiência global, objetivou-se a análise da sua possível relação com variados fatores em nível setorial e organizacional, visando promover uma discussão sobre quais deles se relacionam às melhores práticas observadas na amostra. Sendo assim, primeiramente, verificou-se a relação da eficiência global com os dois fatores utilizados como insumos na análise: quantitativo da equipe e recursos financeiros. Os recursos financeiros e humanos são frequentemente operacionalizados em pesquisas de avaliação de programas, tanto sob perspectiva da eficiência quanto da efetividade. No campo de TD&E, por exemplo, verificou-se que diversos modelos de avaliação como o CIRO, CIPP, IPO, MAIS, BSC, os modelos lógicos e alguns modelos orientados para resultados como o de Kaufman e Keller (1994) preveem expressamente esses recursos – ao lado dos recursos materiais – como condições importantes que devem estar presentes de maneira adequada ou serem planejados e dimensionados para fins de realização das atividades necessárias ao alcance dos objetivos de um programa (Borges-Andrade, 2006a; Bushnell, 1996; Chyung, 2019; Damasceno et al., 2012; Stufflebeam, 2003; Warr, Birdi & Rackham, 1970). A perspectiva da eficiência, no entanto, visa atestar se houve o uso adequado dos recursos, em perspectiva somativa (Suchman, 1967).

Para o presente caso, a relação com a eficiência global visa verificar o nível de disponibilidade desses recursos nos programas de treinamento com melhor desempenho tendo em vista critérios de economicidade e capacidade de promover o maior número de participações em treinamento. As Figuras 39 e 40 apresentam gráficos de dispersão para análise da relação entre as médias desses fatores e dos escores globais para cada organização da amostra.

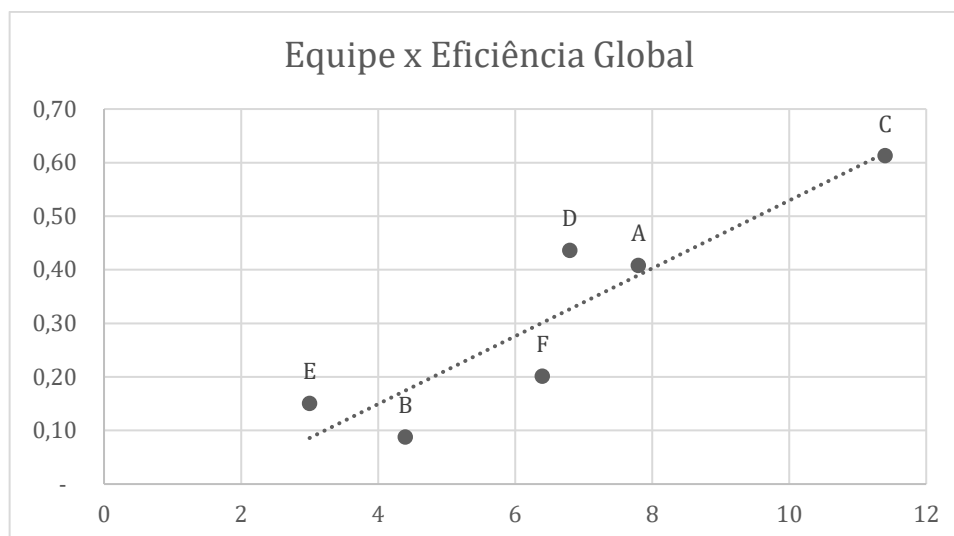


Figura 39. Gráfico Equipe x Eficiência global

Fonte: Elaboração própria.

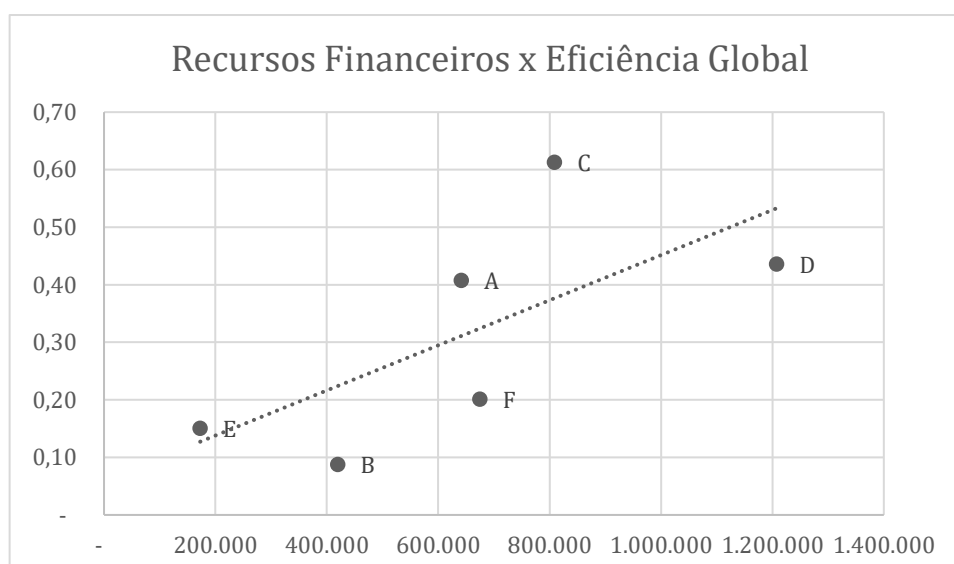


Figura 40. Gráfico Recursos financeiros x Eficiência global

Fonte: Elaboração própria.

Conforme indicam as linhas de tendência das figuras acima, apesar de constituírem os insumos para cálculo do próprio índice de eficiência, uma menor disponibilidade de recursos não resulta em mais eficiência. Ao contrário, quanto maior a quantidade de recursos à disposição da unidade maior a sua eficiência global, uma vez que essa representa não apenas a economia de recursos, mas sim a sua melhor utilização para provisão do máximo de treinamentos para o maior número de servidores.

A breve análise dos gráficos fornece indícios, portanto, de que unidades com mais recursos realizam uma gestão e alocação mais eficiente, podendo indicar a existência de uma escala ótima de equipe e orçamento para uma provisão adequada de treinamentos, que pode ser melhor averiguada em estudos futuros. Essa informação chama a atenção pois, sendo o

investimento financeiro e o quantitativo de servidores da equipe insumos no modelo proposto, poderia ser esperado que unidades com menores quantidades desses recursos fossem mais eficientes em um sentido estrito de eficiência. No entanto, sobretudo a existência de um quantitativo adequado de servidores na unidade se apresenta como um fator relevante para o desempenho da unidade. Tomando como referência as proporções entre a equipe de TD&E e o total de servidores da organização, verifica-se que as unidades mais eficientes globalmente possuem algo próximo a 01 pessoa na equipe para cada 100 servidores na organização, enquanto que nas mais ineficientes há cerca de 01 pessoa na equipe para cada 300 servidores. Mesmo entre organizações com gastos médios similares, observa-se que são consistentemente mais eficientes em termos globais aquelas que possuem maior equipe. O caso da organização “E” em 2018 e 2019, em que se observou ineficiência de escala, demonstra que mesmo com o aumento de investimentos, sem uma equipe adequada, não haverá capacidade para uma execução eficiente desse recurso, o que reforça o indício de que equipes bem dimensionadas de treinamento podem levar a provisões mais eficientes.

Não foram encontradas objeções a esse resultado na literatura, haja vista a previsão de equipes adequadas para a realização adequada de atividades em programas e atividades organizacionais. Porém, a realidade organizacional demonstra que as unidades de treinamento não são compreendidas como estratégicas, sendo frequentemente subdimensionadas. De acordo com o CIPD (2019), as unidades de treinamento tendem a ser as primeiras a sofrerem cortes orçamentários nas organizações. Segundo Camões e Meneses (2016), por sua vez, as unidades de treinamento no setor público federal se veem, frequentemente, com equipes excessivamente reduzidas. Assim, diante de realidades organizacionais de baixa valorização das unidades de treinamento, a realização sistemática de avaliações que demonstrem as consequências negativas de baixo investimento em TD&E – ensejando mais desperdícios e menos resultados – cumpre uma das propostas das metodologias de avaliação de propiciar a conscientização e a criação de novos consensos por meio de dados objetivos (Stufflebeam & Coryn, 2014; Mark et al., 2000). Logo, para o contexto analisado de pesquisa, os resultados comparativos do DEA possibilitam às organizações com menor desempenho a identificação da disponibilidade de recursos como uma contingência a ser gerenciada.

Uma outra contribuição intentada pela presente pesquisa é a análise da relação entre os níveis de eficiência encontrados e as alternativas de investimento em diferentes estratégias instrucionais. A Figura 41 visa estabelecer uma relação entre os níveis médios de eficiência e a proporção gastos com treinamentos divididos com base na modalidade (presencial e a distância) e na organização fornecedora (externa ou interna).

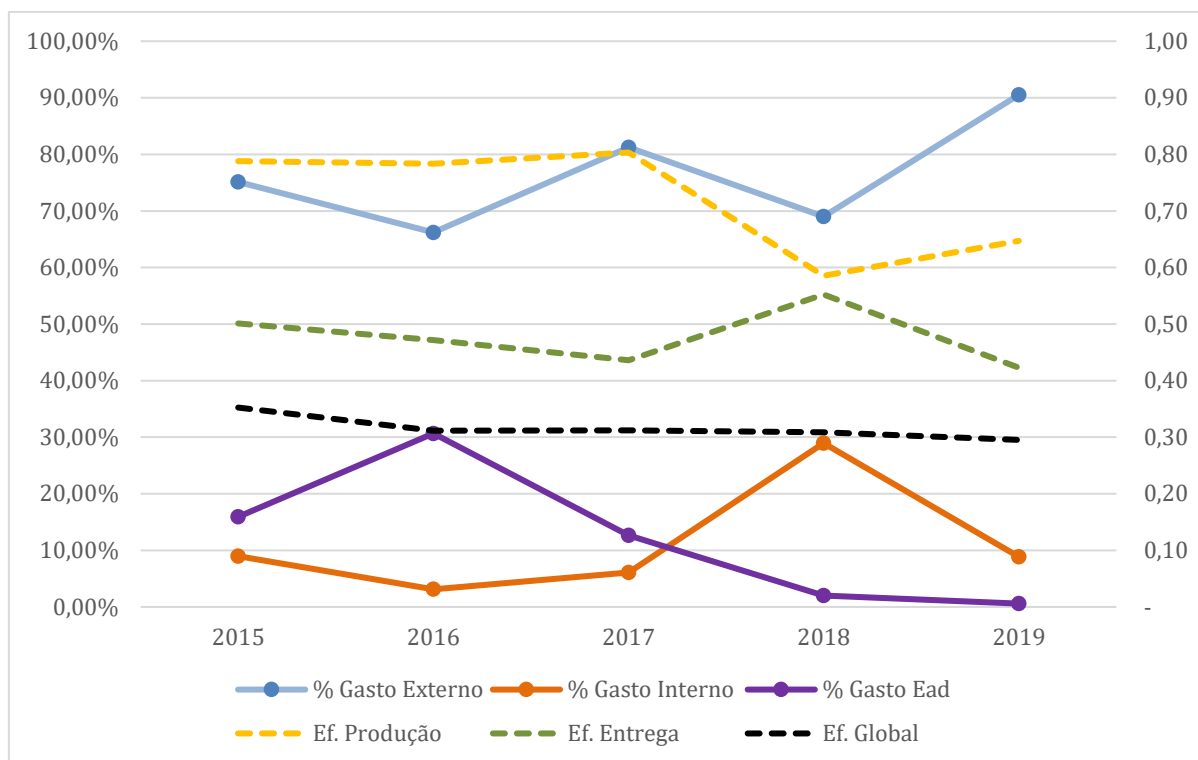


Figura 41. Eficiência e proporção de gastos por modalidade

Fonte: Elaboração própria.

Conforme discutido no início do capítulo de resultados, as organizações da amostra apresentam maior eficiência na contratação de cursos, mas que isso não se reflete, necessariamente, no maior número de oportunidades de capacitação. A orientação para a contratação de cursos no mercado pode ser depreendida da alta proporção de gastos com cursos externos que chega a 90% em 2019. Tal tendência se modifica em 2018 em que há uma melhoria do índice de eficiência de “Entrega”, o qual se associa a um grande aumento nos investimentos para realização de cursos internos. Dessa forma, as diferentes estratégias alternativas de provisão de treinamento, a primeira focada em economia de custos e a segunda focada em maior número de participações, podem se relacionar, na realidade à decisão entre contratar cursos de baixo custo no mercado ou produzir cursos na própria organização.

De forma geral, as organizações da amostra como um todo possuem uma alta proporção de horas de treinamentos realizados por meio de cursos presenciais externos, cuja média relativamente estável em todo o período é de 68%. De acordo com Meneses et al. (2010, p. 22) é prática comum nas organizações recorrer “a treinamentos presenciais, que privilegiam poucas pessoas da organização”, especialmente quando não há mecanismos adequados de planejamento, o que, do ponto de vista dos dados, pode ser uma das razões para os baixos resultados de entrega.

Os cursos internos, por sua vez, são os menos realizados pelas organizações, tendo sido mais executados apenas no ano de 2018. Dessa forma, o aumento na eficiência de entrega naquele ano pode ter se devido ao maior investimento nesse tipo de curso, uma vez que ele depende do planejamento adequado da ação instrucional e visará alcançar grupos de servidores. Essa interpretação é corroborada por Aguinis & Kraiger (2009) que discute evidências de que cursos internos geram maior envolvimento dos funcionários e maior qualidade do treinamento.

O gasto com cursos em modalidade EAD, por sua vez, se reduz progressivamente nas organizações, que se explica pelo exame dos registros nas organizações por uma crescente oferta de cursos gratuitos a distância por Escolas de Governo, como a Escola Nacional de Administração Pública por meio do portal da Escola Virtual de Governo, criada ao final de 2017. Não há uma associação clara dos cursos nessa modalidade com as eficiências médias da amostra, porém, dadas as mudanças para sua aquisição, podem ter sido gerados ganhos de eficiência para as organizações.

Ao analisar-se o nível das DMUs, podem ser tecidas algumas considerações sobre a combinação dessas estratégias. Atendo-se à amostra de organizações estudadas, buscou-se levar em consideração a proporção de horas de treinamento fornecidas com relação à sua modalidade (presencial e a distância) e com relação à organização fornecedora (externa ou interna). O dado proporcional visa demonstrar o quanto cada tipo de capacitação foi privilegiado no âmbito de cada unidade de treinamento e como essas estratégias de implementação se relacionam com a eficiência global. Para capturar essas relações foram elaborados os gráficos de dispersão representados pelas Figuras 42, 43 e 44.

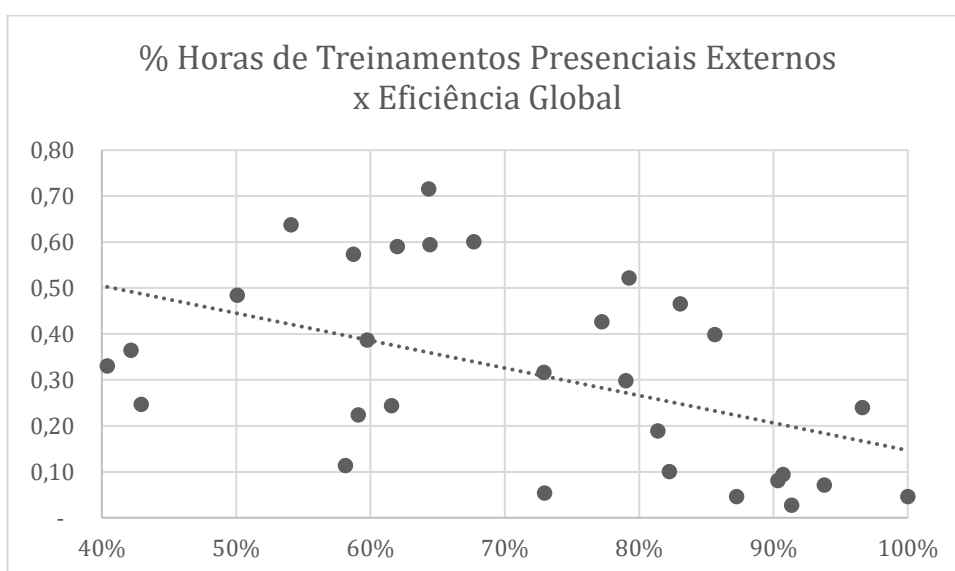


Figura 42. Gráfico Eficiência global x Treinamentos presenciais externos

Fonte: Elaboração própria.

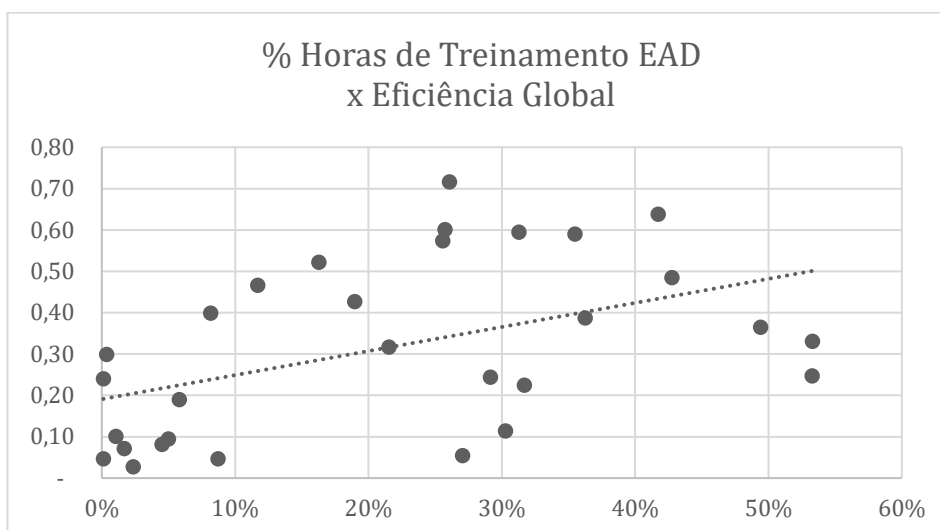


Figura 43. Gráfico Eficiência Global x Treinamento EaD

Fonte: Elaboração própria.

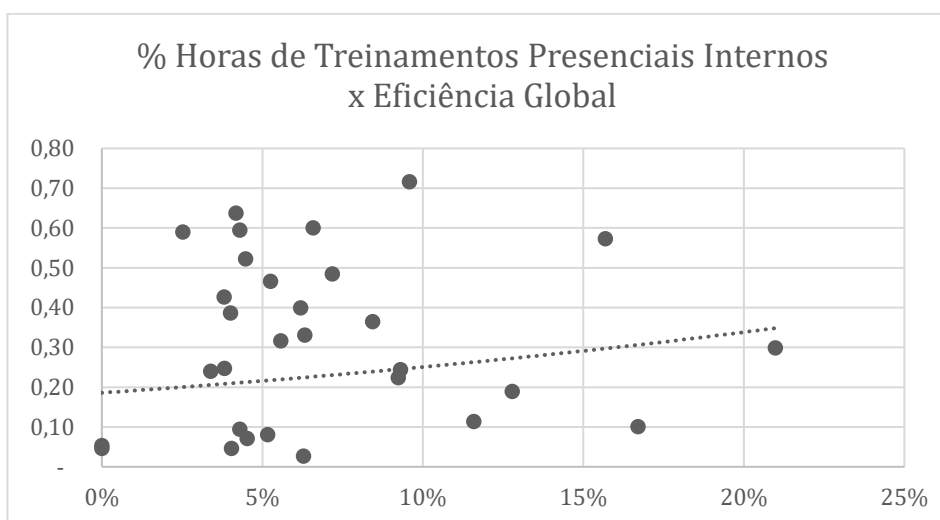


Figura 44. Gráfico Eficiência global x Treinamentos presenciais internos

Fonte: Elaboração própria.

Conforme mencionado, verifica-se a alta proporção de cursos externos, de forma que as organizações foram comparadas a partir desse patamar comum. Verifica-se que as organizações mais eficientes, no âmbito da amostra, realizam entre 50% e 70% das horas de treinamento por meio de cursos dessa natureza, estando em torno da média. Acima desse patamar, uma proporção muito alta de cursos externos começa a se associar a maiores níveis de ineficiência, demonstrando que há relativo desperdício de recursos e redução do alcance das capacitações.

Os cursos à distância (EAD) como alternativa aos presenciais, potencialmente, podem “alcançar um número bem maior de participantes de diferentes unidades da organização” (Meneses et al., 2010, p. 22). Conforme a Figura 43 o EAD, de fato, tende a gerar mais

eficiência, mas não em todos os casos. As organizações mais eficientes possuem entre 25% e 40% da carga horária realizada por meio de EAD, para além disso níveis de ineficiência aumentam, de forma que não é conclusivo que necessariamente promova maior eficiência. No entanto, é mais prejudicial na amostra, a ausência de EAD em que as organizações mais ineficientes possuem menos de 10% da proporção de horas de treinamento nessa modalidade. Por fim, devida à baixa proporção de cursos internos a simples provisão de horas de treinamento não parece guardar uma relação acentuada com eficiência global na amostra, porém nos dados gerais também não havia relação do investimento em cursos internos com a eficiência global, mas positiva com a eficiência de entrega, e negativa com a eficiência de produção.

Dessa forma, as tendências apresentadas corroboram alguns aspectos da literatura de TD&E e constituem indícios que podem ser melhor investigados com a realização de pesquisas correlacionais mais extensivas. Com base nos dados é possível que o aumento da eficiência no Estágio 2 perpassa por uma mudança no perfil das unidades de TD&E investigadas.

Não basta que os profissionais de TD&E sejam bons gestores e garantam contratações eficientes e tempestivas de treinamentos junto a empresas. É preciso que se tenha conhecimento de procedimentos e princípios de tecnologia instrucional para que se consiga planejar programas de treinamento com combinações adequadas de estratégias educacionais, visando maior economia e eficiência, mas também optando por modalidades adaptadas às necessidades do público-alvo. Consequentemente, a provisão de equipes adequadas, tanto em termos de quantidade, quanto do ponto de vista da qualificação e capacitação nos conhecimentos necessários ao planejamento e desenho instrucional é fundamental (Borges-Andrade, 2006a; Fonseca & Meneses, 2016).

Por fim, para além dos aspectos até então apresentados, em geral, sob maior controle das unidades de treinamento, há a possibilidade neste trabalho de suscitar a análise de outros aspectos de caráter contextual. A próxima seção apresenta a análise de outros fatores determinantes cotejados a partir de variáveis do processo de treinamento e da estrutura organizacional.

4.3 Análise de fatores interferentes na eficiência das unidades de treinamento

Para a análise, em caráter exploratório, da influência dos fatores determinantes selecionados como variáveis não-discricionárias, utilizou-se o roteiro proposto por Boueri et al. (2015) com base em Charnes et al. (1981) e que se destina a análise mais simples de variáveis categóricas que permitam a divisão da amostra em subgrupos, os quais são comparados em termos de suas eficiências máximas. Quaisquer diferenças entre os grupos, podem ser associadas aos fatores presentes em um grupo e ausentes no outro.

Tendo em vista que o procedimento iria reduzir a amostra de 30 casos para subgrupos ainda menores complementou-se a análise com duas técnicas aplicadas em outros momentos da pesquisa: o teste não-paramétrico de hipótese de Kolmogorov-Smirnov para comparação das duas subamostras, principalmente em razão de sua adequação a amostras pequenas (Firmino, 2015); e a técnica de *bootstrap*, para reamostragem durante o cálculo da diferença entre as médias dos subgrupos visando estabelecer um intervalo de confiança, representando a diferença mínima e máxima que pode ser observada nos grupos em razão dos fatores analisados. Conforme mencionado no capítulo de Método, se avaliará as possíveis influências da presença (1) ou ausência (0) dos seguintes processos e atividades integrantes do processo de treinamento: Avaliação de Necessidades de Treinamento (ANT); Mapeamento de Competências; Planos Anuais de Capacitação (PAC); Avaliação de Reação; Avaliação de Aprendizagem; Avaliação de Impacto; Planejamento Estratégico com objetivos de capacitação; e Sistema informatizado de capacitação.

Além dessas variáveis, três variáveis relativas ao contexto organizacional foram definidas, mensuradas e dicotomizadas: Porte Organizacional, com organizações com mais (1) e menos (0) de 1.100 servidores; Número de unidades descentralizadas: com organizações classificadas entre aquelas que possuem múltiplas unidades administrativas no território nacional (1) e aquelas que possuem a maior parte dos servidores atuando em uma única localização principal, sede da organização (0); e “Níveis hierárquicos” em que a unidade de treinamento está a até três níveis da cúpula organizacional (1) e aquelas que estão inseridas em estrutura mais abaixo na hierarquia organizacional (0).

Os resultados da análise estão sumarizados para cada fator na Tabela 23, a qual apresenta, primeiramente, os quantitativos de cada subamostra de DMUs, sendo o “Grupo 1” relativo às DMUs com valor “1” no respectivo fator e o Grupo 0 as DMUs com valor “0”. As eficiências médias de ambos os grupos são calculadas representando a eficiência máxima para DMUs de valor “1” (Ef. G1) e a eficiência máxima para DMUs de valor “0” (Ef. G0). As

eficiências são calculadas para o Estágio 1 e para o Estágio 2 com técnica *bootstrap* e comparadas, produzindo uma diferença média entre os subgrupos que visa explicar o efeito do fator sobre a eficiência (Boueri et al., 2015). Para averiguar a significância da diferença média, os subgrupos são comparados por meio do Teste KS que fornece o nível de significância para a hipótese de diferença entre os grupos. São consideradas significantes diferenças com p-valor $> 0,05$ no Teste KS. Por fim, um intervalo de confiança a 95% é calculado para avaliar a extensão mínima e máxima do efeito do fator sobre a eficiência.

Cabe ressaltar que, como pode ser verificado na Tabela 23, alguns fatores possuíam um desequilíbrio grande entre os quantitativos dos subgrupos: Avaliação de Necessidades de Treinamento (26 x 4), Plano Anual de Capacitação (27 x 3), Avaliação de Reação (25 x 5), Avaliação de Aprendizagem (6 x 24), Avaliação de impacto (3 x 27) e Sistema Informatizado (5 x 25). Com isso, as análises acabam por comparar, em geral, uma única organização com todas as outras, não havendo, portanto, uma variação que permita extrair conclusões válidas e adequadas nestes casos. Sendo assim, serão aprofundadas as discussões apenas dos fatores que possuem uma distribuição relativamente equilibrada de DMUs nos Grupos 1 e 0. Tais fatores estão sombreados na Tabela 23 e serão comentados sequencialmente com relação às diferenças significativas observadas em cada um dos estágios e se possuem influência positiva (Pos.), negativa (Neg.) ou indeterminada (Ind.).

Tabela 23
Efeitos das variáveis não-discrecionárias nas eficiências

Fator	G1	G0	Estágio 01							Estágio 02						
			Efic. G1	Efic. G0	Diferença Média	Mín	Máx	Teste KS	Efeito	Efic. G1	Efic. G0	Diferença Média	Mín	Máx	Teste KS	Efeito
ANT	26	4	100%	85%	15%	3%	29%	0,04	Pos.	90%	47%	44%	8%	69%	0,08	-
GPC	12	18	98%	91%	7%	1%	12%	0,01	Pos.	89%	95%	-6%	-15%	3%	0,40	-
PAC	27	3	96%	100%	-4%	-6%	-2%	0,30	-	100%	18%	82%	78%	89%	0,01	Pos.
Av. Reação	25	5	98%	85%	13%	-3%	37%	0,52	-	100%	19%	81%	78%	87%	0,00	Pos.
Av. Aprendizagem	6	24	78%	100%	-22%	-47%	-5%	0,00	Neg.	52%	100%	-48%	-52%	-45%	0,00	Neg.
Av. Impacto	3	27	67%	100%	-33%	-83%	0%	0,18	-	51%	100%	-49%	-58%	-43%	0,01	Neg.
Plan. Estratégico	20	10	95%	92%	3%	-5%	9%	0,07	-	85%	97%	-12%	-22%	-3%	0,24	-
Sist. Informatizado	5	25	91%	100%	-9%	-15%	-4%	0,01	Neg.	50%	100%	-50%	-56%	-45%	0,00	Neg.
Regionais	15	15	94%	81%	13%	0%	26%	0,66	-	59%	95%	-36%	-46%	-25%	0,00	Neg.
Posição hierárquica	17	13	90%	91%	0%	-12%	9%	0,03	Ind.	87%	81%	6%	-9%	23%	0,02	Neg.
Porte	15	15	85%	99%	-14%	-22%	-8%	0,00	Neg.	85%	85%	0%	-14%	14%	0,00	Ind.

Nota. Fonte: Elaboração própria.

Com as devidas ressalvas, visto que se trata de uma comparação de 26 DMUs contra 4 apenas, a ANT apresentou alguma contribuição para o aumento da eficiência de produção em cerca de 15%. Tais resultados demonstram, portanto que esses processos podem ser uma ferramenta relevante para o planejamento da capacitação de forma a proporcionar, como no caso da amostra investigada, o uso mais racional dos recursos pelas organizações, corroborando prescrições da área de TD&E (Borges-Andrade, 2006a; Ferreira & Abbad, 2008; Meneses et al., 2010; Salas et al., 2012).

Assim, a ANT que vem sendo retratada na literatura como fator fundamental para a efetividade do treinamento (Alvarez et al., 2004), pode apresentar ganhos também de eficácia. Entretanto, chama a atenção o fato de que o processo ainda vem sendo realizado de maneira incipiente pelas organizações, segundo o que foi verificado nas entrevistas preliminares: embora a maioria das organizações tenha relatado que executa a ANT, ela ocorre majoritariamente via preenchimento de formulários, sem a adoção de procedimentos sistemáticos que visem verificar a existência de lacunas de competências. Tal fato, é corroborado pela verificação de que apenas 12 DMUs possuem Mapeamento de Competências (GPC), diante de 26 DMUs que realizam ANT.

Tratando-se especificamente sobre a variável GPC, isto é, a presença de Mapeamento de Competências, foi relatado como um processo realizado por 12 DMUs as quais apresentaram, para o Estágio 01, uma eficiência em média 7% superior às DMUs que não a utilizam. Nesse caso, o efeito positivo (Pos.) é confirmado pelo cálculo dos intervalos de confiança que preveem uma variação positiva entre 1% e 12% do escore de eficiência nas unidades que utilizam essa técnica de gestão da capacitação. Nesse sentido, há evidências de que a gestão por competências pode promover maior eficiência sobretudo na etapa de aquisição e produção de treinamentos, ao orientar as necessidades de aprendizagem.

O segundo fator abordado na análise e de grande relevância em toda a literatura refere-se à existência de um plano estratégico, que defina objetivos de capacitação, como forma de efetividade da função de treinamento (Meneses et al., 2010; Freitas & Mourão, 2012; OCDE, 2010), sendo também um dos critérios previstos pela PNDP (Brasil, 2019a). Para a presente pesquisa questionou-se aos respondentes nas organizações pesquisadas acerca da existência de planejamentos estratégicos formalizados e foi realizada a análise documental dos respectivos planos para averiguação da relação com as atividades de capacitação. Nesse ínterim, 20 DMUs foram identificadas com um plano estratégico organizacional formal no qual constavam objetivos estratégicos de capacitação. No entanto, os testes não foram significativos para ambos os Estágios, não permitindo concluir que tal fator tenha efeito sobre a eficiência observada nas

DMUs. Além da não-significância do teste, o intervalo de confiança mostrou-se inconclusivo posto que varia em torno de “0”.

Um terceiro fator analisado, de natureza estrutural no nível organizacional, refere-se à sua distribuição em diferentes estados da federação, ensejando um maior nível de complexidade às unidades de treinamento que devem ser capazes de capacitar servidores em todo o território nacional. A amostra divide-se igualmente entre organizações e respectivas DMUs com unidades altamente descentralizadas (Grupo 1) e aquelas que operam majoritariamente em uma única localidade (Grupo 0). Não foi encontrada relação com o Estágio 01 de produção, mas verificou-se uma diferença significativa e que representa um impacto considerável nos índices de eficiência para o Estágio 02, de entrega do treinamento, refletindo-se em prejuízos para o alcance das ações educacionais. Uma organização descentralizada pode imputar uma perda de eficiência de entrega que varia entre 25% e 46%, sendo observada uma perda média de 36%, ou seja, uma perda de 36% da capacidade de alcançar uma maior proporção de servidores da força de trabalho. Ressalta-se que esse fator foi encontrado nas organizações “A”, “B” e “E”, de forma que essa pode ser uma das razões que justificam o baixo desempenho comparativo na Etapa 02. Nesse sentido, uma alternativa para essas organizações é a descentralização da função de treinamento com a criação de subsidiárias regionais com capacidade de implementar programas de treinamento localmente, ampliando o alcance da unidade.

Um outro fator estrutural refere-se ao porte da organização que, de forma similar à descentralização, enseja um maior desafio à unidade de treinamento que opere em organizações com uma força de trabalho muito grande. A análise indicou uma diferença significativa entre organizações com mais e menos de 1.100 servidores (aproximadamente a mediana da amostra). As unidades de treinamento em organizações de grande porte tendem a ter menor eficiência no Estágio 01, com perda média de 14% frente a unidades em organizações menores, sendo esse efeito negativo confirmado pelo intervalo de confiança que estabelece uma perda entre 8% e 22% de eficiência. Para o Estágio 02 a influência desse fator apresenta-se como inconclusiva, sendo em média, uma influência nula.

Por fim, o fator posição hierárquica visa refletir a autonomia estrutural e a proximidade da DMU junto às instâncias estratégicas da organização, facilitando a obtenção de suporte e recursos para suas atividades (Côrtes & Meneses, 2019; Fonseca et al., 2013). As DMUs com altas posições hierárquicas foram classificadas no Grupo 1 e comparadas com aquelas mais abaixo na hierarquia organizacional, classificadas no Grupo 0. Esse fator apresentou-se como inconclusivo com relação a sua influência sobre os níveis de eficiência observados. Houve uma pequena tendência de perda de eficiência de produção (Estágio 01) para o Grupo 1, mas uma

tendência de incremento positivo em 6% na eficiência de entrega (Estágio 02), o que pode ser reflexo da maior autoridade e autonomia auferida pela unidade para promoção e divulgação de políticas e programas de treinamento em toda a organização, um aspecto político observado por Borges-Andrade (2006b).

A análise realizada, portanto, distingue-se da análise mais tradicional de aplicação de métodos estatísticos para realização de regressão (Boueri et al., 2015; Ferreira & Gomes, 2020), porém apresenta como vantagem a maior simplicidade de mensuração e análise para que sejam exploradas possíveis influências ambientais sobre os escores de eficiência das organizações. Cabe ressaltar que o objetivo ao realizar este tipo de análise não reside em obter conclusões generalizáveis, mas sim identificar tendências que possam guiar novas práticas e estudos. Para o presente caso, por exemplo, observa-se que as organizações “C” e “D”, que apresentaram escores de eficiência consistentemente maiores que as demais organizações, operam em condições mais favoráveis. São unidades que 1) ocupam uma posição hierárquica elevada; 2) utilizam ANT e Mapeamento de Competências; 3) atuam em organizações com menos de 1.100 funcionários; 4) estão localizadas em uma sede principal única; e 5) possuem maior quantitativo de servidores na unidade de capacitação. Dessa forma, é possível que, não estando presentes esses fatores, tais DMUs pudessem não apresentar índices de eficiência tão elevados. A análise também acrescenta elementos presentes nos *benchmarks* que talvez não possam ser plenamente copiados pelas demais organizações, haja vista serem fatores não-discrecionários.

A discussão sobre esse tipo de fator reforça a importância de se conhecer o ambiente de atuação das unidades de treinamento e tecer estratégias adequadas para se minimizar ou modificar influências contextuais. Entretanto, no caso da amostra investigada, a análise permaneceu inconclusiva ou pouco válida para a maior parte das variáveis relativas aos processos de treinamento. Não obstante, os achados acerca da influência positiva ou neutra, respectivamente com relação ao mapeamento de competências e o alinhamento à estratégia organizacional, terminam por contribuir para apoiar a gestão de unidades de treinamento, evidenciando quais aspectos são relevantes e a dimensão de seus ganhos para a eficiência para que sejam implementados e aprimorados por gestores e servidores da área de TD&E. Visando congrega as implicações práticas para as organizações estudadas, a próxima seção apresentará as recomendações que podem ser apresentadas às organizações tendo em vista a interpretação conjunta dos dados elencados.

4.4 Recomendações e implicações para as organizações

A avaliação de eficiência tem como objetivo apontar possíveis melhorias aos programas e atividades organizacionais para a geração de mais resultados tendo em visto os recursos disponíveis. A eficiência é um processo complexo e multideterminado, aspecto que se comprova na presente análise em que as eficiências das organizações e as recomendações decorrentes se relacionam a questões de natureza microeconômica como a eficiência técnica e de escala, questões relativas aos processos e estratégias de implementação do TD&E e, por fim, questões de natureza administrativa relacionadas ao contexto do setor público. As recomendações para cada uma das organizações, portanto, se fundamentam nos resultados de pesquisa obtidos para apresentar explicações para os seus desempenhos com base nessa perspectiva interdisciplinar, tecendo considerações sobre as possíveis melhorias que podem ser implementadas.

Conforme observado, há evidências de que a realização excessiva de cursos externos ainda que mais econômica, não promove o maior alcance das ações educacionais. A produção de cursos internos, por outro lado tem efeito contrário em que, há um aumento de gastos, mas que promove um maior número de participações. Como evidência desse aspecto, a maior parte das organizações possui ineficiências de escala no Estágio 02 implicando uma relação desvantajosa entre a quantidade de horas fornecidas e a quantidade de participações, acentuada principalmente nas organizações que contratam apenas vagas avulsas e promovem ou planejam poucos cursos em turmas. Como segunda recomendação geral, as unidades não podem alcançar índices de eficiência adequados sem que possuam equipes bem dimensionadas e preparadas para implementação de um processo de treinamento sistemático, alinhado a princípios de planejamento instrucional (Borges-Andrade, 2006b) e, caso possível, baseado em competências. Conforme se observará na amostra, os desempenhos superiores das organizações advêm de processos bem estruturados de TD&E, equipes bem dimensionadas e decisões adequadas sobre estratégias de treinamento que reflitam a capacidade de utilizar diferentes alternativas instrucionais, com uma combinação de tipos e modalidades de treinamento.

Realizadas essas duas recomendações gerais, as demais seções buscaram realizar uma descrição dos casos e fornecer, com base nos dados de pesquisa, recomendações mais específicas aos casos de cada organização. Para tanto, é apresentado um painel que visa agregar os principais indicadores considerados relevantes para a análise.

4.4.1 Organização A

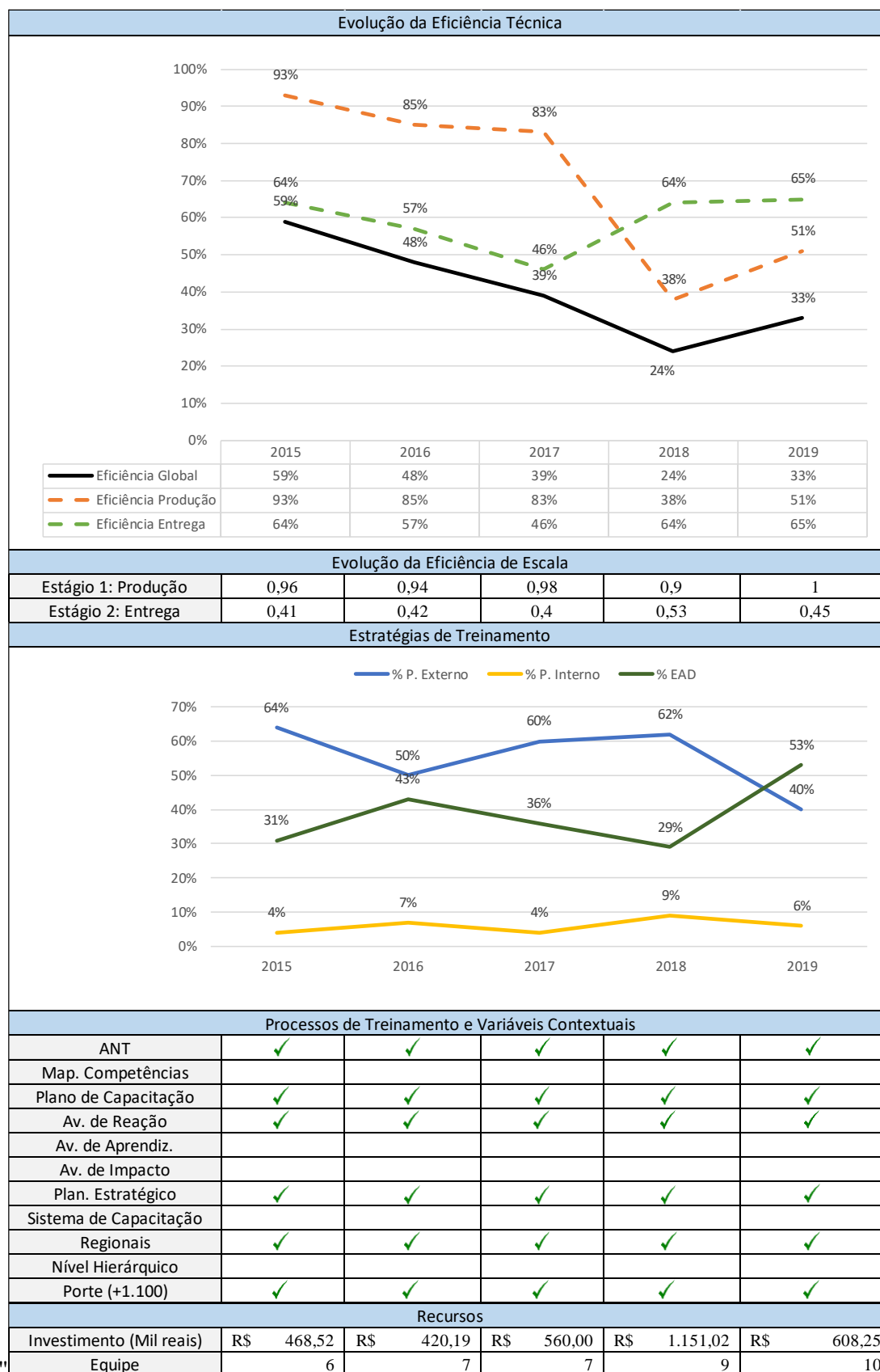


Figura 45. Eficiência e Variáveis contextuais da Organização A

Fonte: Elaboração própria.

A organização “A” foi uma das que apresentou um desempenho mediano em comparação às demais organizações da amostra, podendo ser considerada, não obstante, a organização com o terceiro melhor desempenho, em geral, antecedida pelas organizações “C” e “D”. “A” se caracteriza por ser uma organização de grande porte e que possui uma grande quantidade de unidades regionais e postos de fiscalização distribuídos em todo o país. Suas características estruturais lhe impõem, portanto, uma desvantagem para a provisão eficiente de cursos, visto que esses dois aspectos se relacionam a uma queda média de 15% de eficiência de produção e 36% para entrega. Tais fatores, não são observados em “C” e “D” o que pode explicar o posicionamento de “A” em relação às demais.

O processo e as estratégias para implementação do treinamento realizado por “A” podem ser considerados similares aos *benchmarks* da amostra, em que há o levantamento de necessidades de treinamento, o qual, porém, é realizado por meio de formulários, e o fornecimento dos cursos é majoritariamente feito com contratações ou parcerias junto a fornecedores externos. Não obstante, a análise documental relativa aos planos de capacitação de “A” demonstra que há um processo estruturado de planejamento para um período bianual, com a construção de um programa de treinamentos com clara descrição dos cursos a serem realizados, divididos em áreas temáticas, unidades organizacionais e regionais atendidas. Os cursos são associados sistematicamente a objetivos instrucionais, resultados esperados, alternativas de implementação, assim como às estratégias e metas organizacionais. Além disso, “A” adota metas de capacitação (% de capacitados) previstas no plano estratégico da organização, de forma que as atividades são acompanhadas pela alta administração. Por fim “A” utiliza de forma sistemática avaliações de reação para a maior parte dos cursos e solicita relatórios para acompanhamento dos resultados da capacitação. Com relação ao mapeamento de competências, relatou-se que os mesmos realizaram um piloto e pretendem implementar o modelo futuramente. Ressalte-se que o desenvolvimento desse processo estruturado de capacitação foi acompanhado do incremento da equipe de 06 para 10 pessoas, todas com formação superior e algumas com especialização.

Para o caso de “A”, portanto, as recomendações que podem ser suscitadas para o seu processo de treinamento residem na realização do mapeamento de competências, com vistas a aprimorar ainda mais o seu processo de planejamento. Sob o ponto de vista da eficiência, o mapeamento de competências está associado a um melhor desempenho no Estágio de “Produção”. Além disso, aproveitando a vinculação estratégica dos treinamentos com metas organizacionais, e a cultura de relatórios pós-treinamento, “A” pode buscar operacionalizar avaliações de aprendizagem e impacto para aprimoramento de ações educacionais.

Sob o ponto de vista da eficiência a adoção de estratégias de EAD em 2019 se relacionou à recuperação do índice de eficiência global que vinha se deteriorando, o que pode estar relacionado ao melhor alcance de servidores em unidades descentralizadas. Assim, ainda que na amostra a adoção do EAD tenha se mostrado inconclusiva como estratégia de melhoria da eficiência, para “A” parece ter vantagens em razão de suas características estruturais, podendo ser feita uma recomendação no sentido de maiores investimentos nessa modalidade. Conforme Meneses et al. (2010, p. 22) os treinamentos à distância têm como vantagem “alcançar um número bem maior de participantes de diferentes unidades da organização”.

Por fim, a análise de *benchmarks* demonstrou que “A” deveria ter como referência (*peer*) a DMU “C7”. No entanto, haja vista as características estruturais de “A” pode não ser possível a replicação das estratégias daquela DMU, sendo mais indicada o referenciamento em “F6”. A estratégia de F6 consistiu na promoção de cursos internos para grandes quantidades de servidores o que, para “A” pode aperfeiçoar sua eficiência de escala de Entrega aumentando seu desempenho.

4.4.2 Organização B

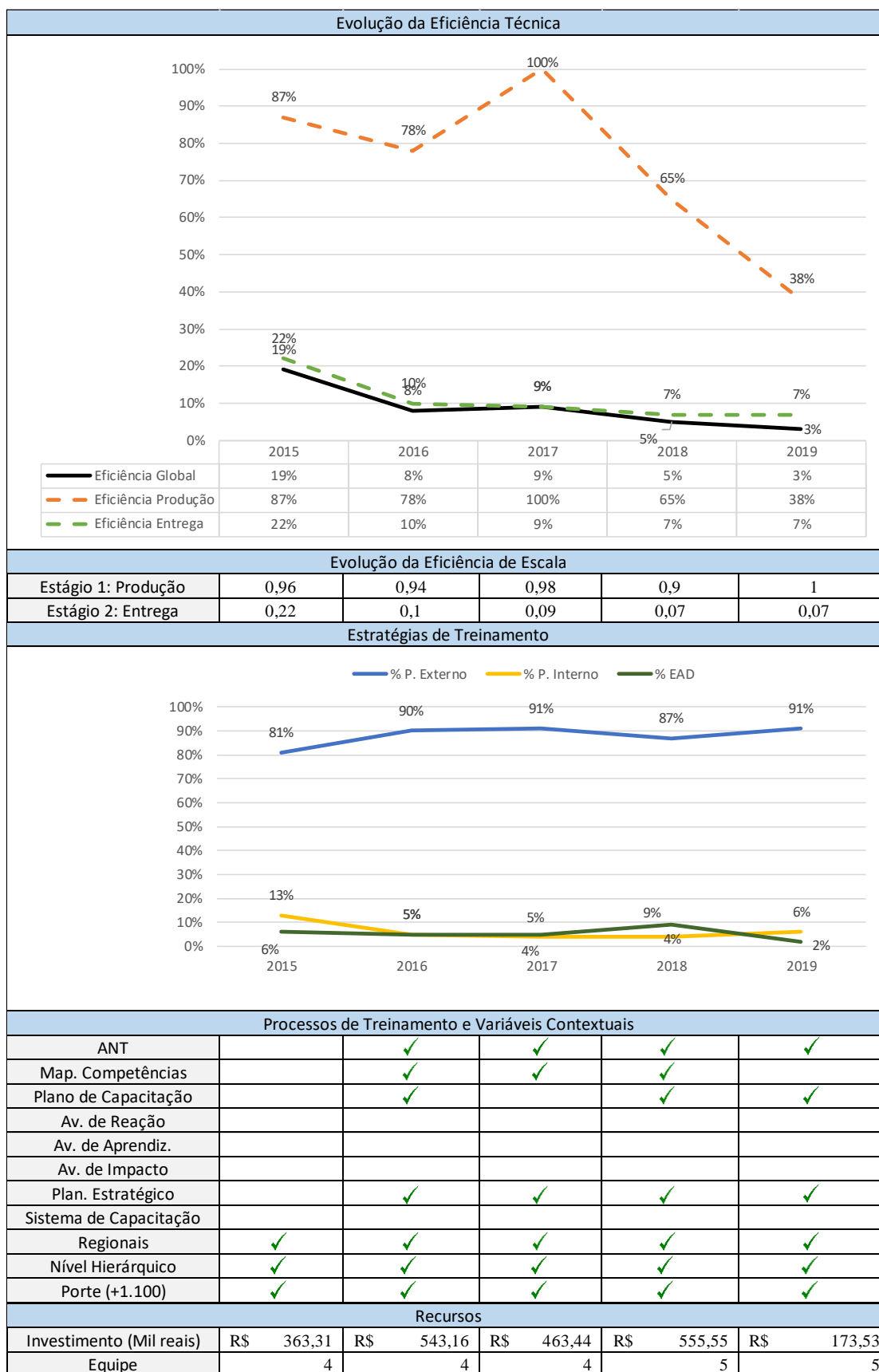


Figura 46. Eficiência e Variáveis contextuais da Organização B

Fonte: Elaboração própria.

A organização “B” conta com as mesmas características estruturais de “A”, sendo uma organização de grande porte distribuída em unidades regionais e postos de fiscalização em todo o país o que leva, portanto, às mesmas desvantagens à eficiência de produção e entrega de ações educacionais. No entanto, “B” não conta com os mesmos recursos e maturidade de processos de planejamento e implementação de treinamentos que “A”. O volume de recursos financeiros de “B” sugere uma situação de redução do investimento nessa atividade e o tamanho de sua equipe é bastante reduzido. Conseqüentemente, “B” terminou sendo posicionada dentre as DMUs com pior desempenho na amostra.

A análise dos índices (Figura 46) deixa evidente algo que se comprova nos dados de capacitação de “B” que é a adoção de uma estratégia fortemente orientada ao fornecimento de vagas avulsas de baixo custo em cursos e eventos externos, agravado por ser um quantitativo muito baixo de ações para poucos indivíduos e, em muitos casos, acerca de temas administrativos, não relacionados às atividades finalísticas da organização. Os reflexos dessa estratégia se verificam na ineficiência de escala de entrega. O melhor desempenho de “B” ocorreu justamente quando houve um aumento no fornecimento de cursos internos em 2015 sendo implementados, conforme os dados, programas gerenciais, cursos de formação e cursos de longa duração em temas administrativos e finalísticos para múltiplas turmas de servidores.

O desempenho de “B” reflete o seu processo assistemático de planejamento, o qual não ocorre de maneira consistente em todo o período. Observa-se na análise documental que o planejamento em B6 apresenta pouco detalhamento das ações a serem executadas e B8 consistiu em um catálogo de cursos fornecidos por instituições externas. Conforme relatos dos gestores, “B” busca realizar a ANT por meio do envio de planilhas às unidades e que buscou a implementação de competências mapeadas como ferramenta de planejamento, mas que se verificou que estas estavam desatualizadas. Conforme relatado por “B”: “muitas unidades organizacionais tiveram dificuldades em vincular as capacitações às competências mapeadas.”

Dessa forma, de maneira geral, a principal recomendação para o aperfeiçoamento do desempenho de “B” relaciona-se à implementação de processos efetivos de tecnologia instrucional, sobretudo de planejamento do programa de treinamentos de maneira sistemática (Borges-Andrade, 2006b; Meneses et al., 2010). Como ressalva, a análise documental demonstra que “B” tem priorizado o investimento em programas de pós-graduação e outros tipos de ações, o que pode remeter a objetivos e prioridades distintas oriundas de diretrizes estratégicas, da cultura ou da natureza das atividades organizacionais. Por fim, a análise de *benchmarks* demonstrou que “B” deve ter como referência (*peer*) a DMU “F6”, similar ao caso de “A” comentado anteriormente.

4.4.3 Organização C

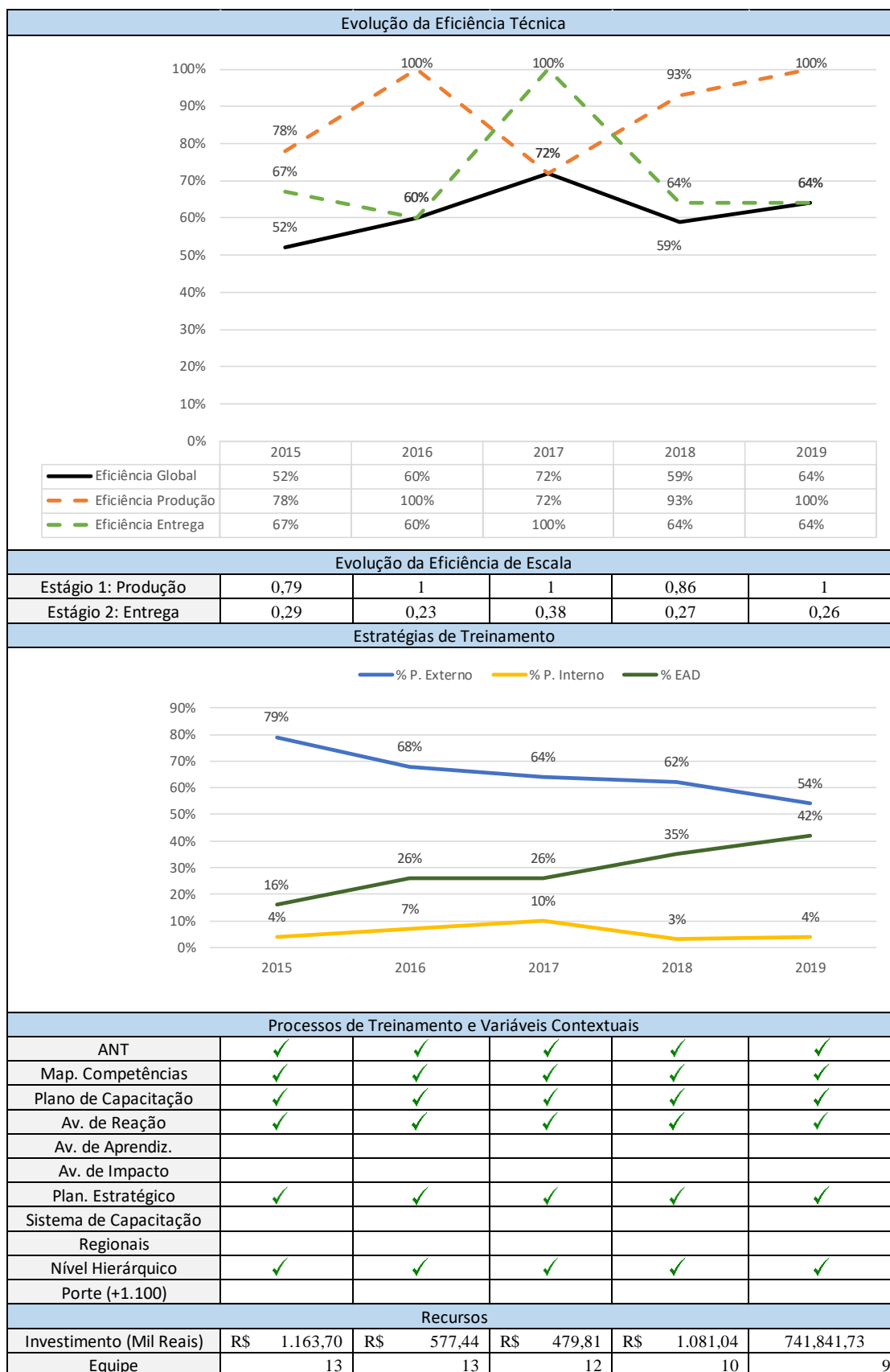


Figura 47. Eficiência e Variáveis contextuais da Organização C

Fonte: Elaboração própria.

A organização “C” apresenta-se como a que possui melhor desempenho global em toda a amostra, sendo os melhores as DMUs C6, C7 e C9. Conforme discutido na seção anterior, “C” atua em condições que se mostraram favoráveis à maior eficiência, sendo uma organização de menor porte, localizada em uma sede principal única e bom posicionamento hierárquico, denotando maior autonomia gerencial. Somado a isso, “C” apresenta um dos processos mais estruturados e padronizados de planejamento.

A análise dos planos de capacitação de “C” demonstra que em C5 foi desenvolvida uma metodologia de ANT e de planejamento que foi replicada e aperfeiçoada anualmente. A ANT se baseia na consulta a gestores formalmente responsáveis pela análise e sistematização das necessidades de desenvolvimento em suas unidades organizacionais. Dessa forma a ANT é descentralizada, com atribuição de cotas orçamentárias aos gestores responsáveis nas unidades, os quais devem indicar os cursos a serem realizados identificando os objetivos estratégicos a que se vinculam, as questões específicas de desempenho que se propõem a resolver e as perspectivas de longo prazo, estabelecendo uma trilha de aprendizagem. As solicitações são analisadas e agregadas pela unidade de treinamento subsidiando um plano que detalha os principais cursos que deverão ser implementadas pela própria unidade – internos – e as previsões dos demais cursos considerados estratégicos. Nesse período “C” manteve metas quantitativas e de carga horária de cursos, assim como orçamentárias e de investimento por servidor, sendo que parte desses indicadores integra o planejamento estratégico organizacional, sendo acompanhado pela alta administração.

A partir de C7 as competências que haviam sido mapeadas anteriormente são aperfeiçoadas e incluídas no âmbito do planejamento servindo de base para definição de trilhas de aprendizagem com definição dos treinamentos a serem realizados para o desenvolvimento de cada uma das competências divididas em categorias técnicas, gerenciais e específicas. “C” mantém cerca de quatro programas permanentes realizados ciclicamente em todos os períodos, dentre eles os treinamentos regulares e um programa permanente de desenvolvimento gerencial. Parte significativa dos cursos promovidos por “C” é realizado por meio de parcerias, por vezes com menor custo, o que permite a manutenção de uma oferta de regular junto a instituições parceiras. Por fim, “C” realiza sistematicamente a avaliação de reação por meio de formulários eletrônicos enviados após a conclusão dos cursos.

As equipes de “C” têm sido reduzidas ao longo do período, o que pode prejudicar a sua capacidade de realização das atividades de seu processo de TD&E. Não obstante, seu quantitativo se manteve elevado e, de forma similar a “A”, é constituída de pessoal com nível superior e parte deles com especialização.

O caso de “C” evidencia que os índices de eficiência global foram capazes de capturar as características latentes de maturidade de processos e competência gerencial que se expressaram nos resultados de desempenho. Em outras palavras, o DEA capturou a melhor tecnologia da amostra. Dessa forma, poucas recomendações podem ser feitas a “C” com base nos dados, haja vista que esta é o principal *benchmark* para as demais. No entanto, o *benchmark* para “C” é seu próprio desempenho em C7, o melhor da amostra, com a maior eficiência de entrega, obtida em um período em que realizou uma grande quantidade de cursos internos dentre eles um programa de cursos estratégicos envolvendo todos os servidores da organização. A eficiência de entrega de “C”, portanto, acompanha o indicador de cursos internos e há uma redução dos cursos externos em favor dos cursos EAD a partir de C7 o que promoveu a melhoria da eficiência de produção indicando que houve melhor custo benefício.

A ineficiência de escala de entrega sugere que a realização de cursos para muitas pessoas não é a principal estratégia de “C”, mas a oferta de uma grande quantidade e variedade de cursos mesmo com poucas participações. Considerando o total de treinamentos de toda a amostra, “C” foi sozinha, responsável em média por 35% de todos os cursos ofertados ao longo dos períodos por todas as organizações. Com essa estratégia, alcançou altos índices de eficiência, apesar das ineficiências de escala.

4.4.4 Organização D

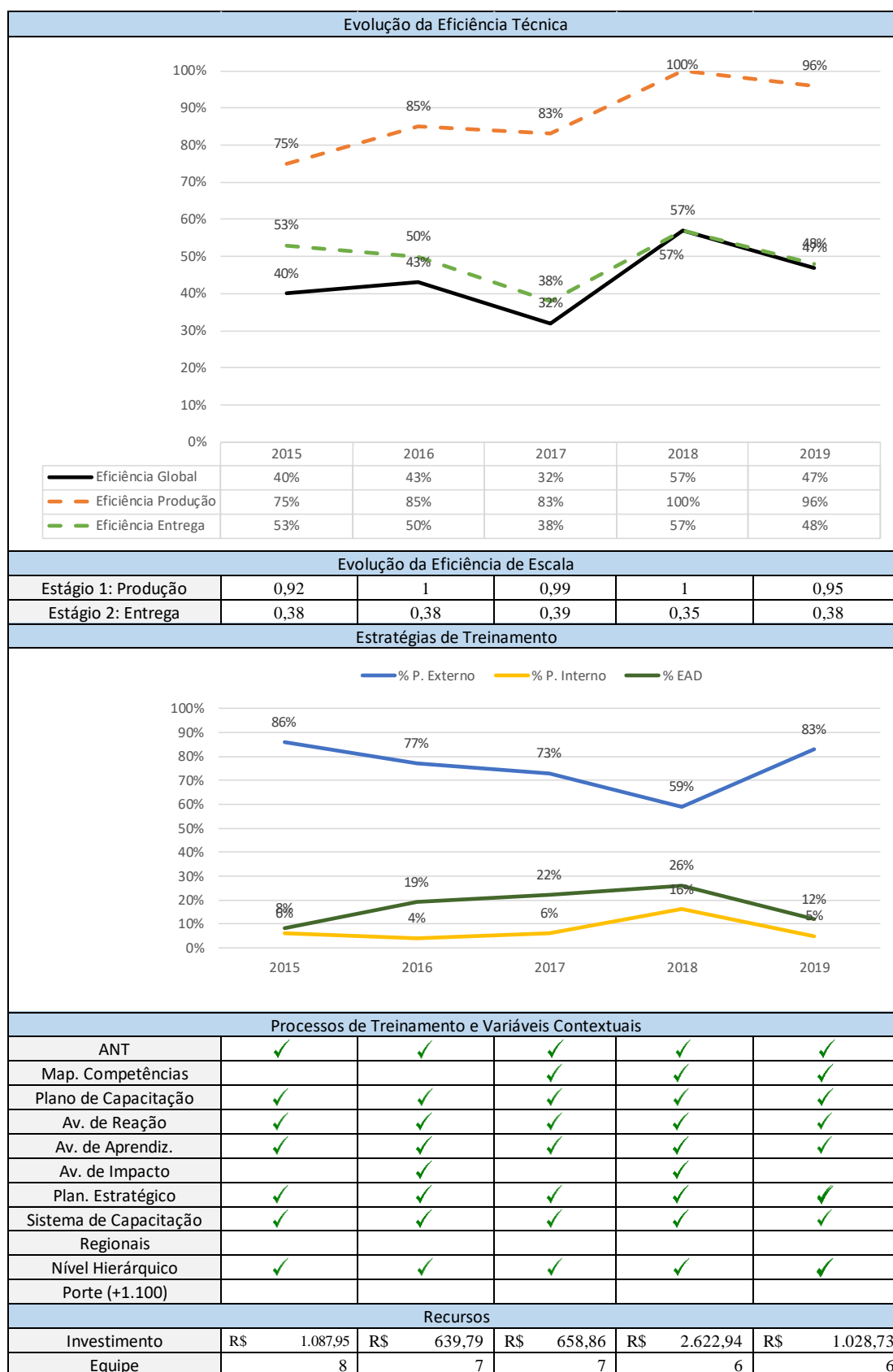


Figura 48. Eficiência e Variáveis contextuais da Organização D

Fonte: Elaboração própria.

A organização “D” também se apresenta como uma das melhores da amostra e possui as mesmas características estruturais de “C” mais vantajosas para a eficiência do treinamento. A equipe de “D” tem o maior nível de formação com todos os integrantes com pós-graduação. Seus níveis de eficiência global, no entanto se mantém intermediários, acompanhando a eficiência de entrega que, por sua vez, acompanha, novamente, os cursos internos. Cabe notar que, mesmo com a redução dos cursos externos que seriam os “mais baratos”, D8 foi eficiente com relação à produção, uma tendência oposta à de outras organizações.

Os processos de planejamento de “D” podem ser considerados similares ao de “C”. A mesma metodologia descentralizada de ANT é utilizada e subsidiam um plano em que são descritos os cursos que compõem programas de treinamento específicos para áreas-chave da organização. Os treinamentos não são plenamente detalhados nos planos, mas são associados a um modelo de competências baseado em descrições comportamentais, desdobrado em níveis e cargos. Os cursos são classificados em termos de complexidade e em perspectiva técnica e gerencial. “D” realiza a gestão de treinamento por meio de sistemas eletrônicos, os quais também apoiam a realização de avaliação de reação para todos os cursos e de aprendizagem e impacto para cursos considerados estratégicos. De forma geral, “D” possui a mesma tecnologia, estratégia e condições de atuação de “C”, que também lhe é o principal *benchmark*. Sendo tão similares as recomendações a “C” se aplicam também a “D”. Possivelmente os melhores desempenhos de “C” advenham de características do processo mais consolidado de planejamento do treinamento, maior equipe ou decisões gerenciais, que podem ser tomadas como possíveis recomendações mais simples a “D”.

4.4.5 Organização E

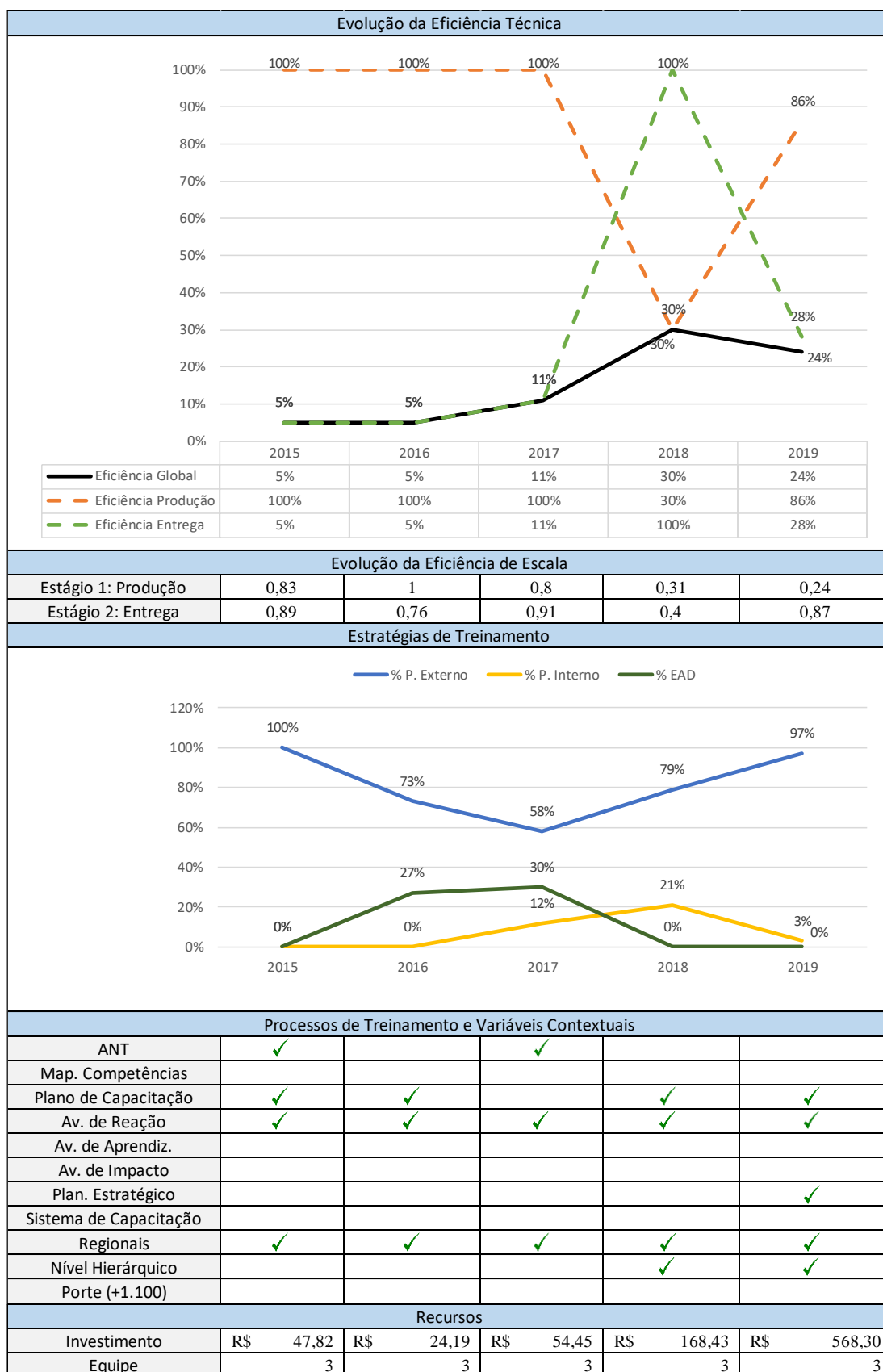


Figura 49. Eficiência e Variáveis contextuais da Organização E

Fonte: Elaboração própria.

A organização “E” caracteriza-se por ser uma organização de médio porte com um certo nível de descentralização, possuindo unidades regionais distribuídas no país. O posicionamento de “E” nos diagramas do DEA demonstravam que sua unidade de treinamento operava em pequena escala, com equipe bem reduzida, responsável pela execução de valores muito pequenos para a aquisição de cursos de baixo custo para um pequeno grupo de servidores. Esse perfil, que vigora nos períodos E5, E6 e E7 explica a sua eficiência econômica e de escala da unidade e sua ineficiência de entrega. Nesse período a unidade não possuía um processo de planejamento e realizava em média, cerca de 30 treinamentos ao ano, 14 vezes menos do que “C”. No entanto “E” possui um maior número de servidores que “C”.

A situação de “E” se altera com a mudança de governo, quando passa por uma reestruturação organizacional e mudança de gestão em que a unidade de treinamento passa a ocupar uma posição mais elevada na hierarquia. Dessa forma “E” é uma unidade em transição de um pequeno departamento de treinamento para uma unidade com maior poder e mais recursos, o que indica que a nova estrutura e a liderança organizacional compreende o seu papel como mais estratégico.

A partir de E8 os valores para treinamento aumentam consideravelmente, chegando ao maior patamar em E9. Em E8 se observa uma mudança de estratégia de fornecimento nos treinamentos que promoveu mudanças nos índices de eficiência. De acordo com os registros de treinamento, “E” promoveu uma série de programas internos de cursos técnicos e de desenvolvimento gerencial de longa duração, tornando-a eficiente no quesito entrega. No entanto, a unidade manteve o mesmo quantitativo de equipe – três pessoas com nível superior – o que pode ter sido um obstáculo para a implementação de um processo mais estruturado de planejamento do TD&E. Dessa forma, tendeu a manter em E9 seu padrão de priorizar a contratação de cursos de mercado, ainda que tenha mantido alguns cursos internos de desenvolvimento gerencial. As recomendações principais para “E” seriam as mesmas para “B”, com a diferença que “E” encontra-se em um momento favorável para buscar um papel mais estratégico na organização e tem criado uma nova estratégia, que deve ser ampliada, de promover cursos mais alinhados às necessidades organizacionais. Porém, ainda precisará desenvolver uma equipe capaz de implementar um processo efetivo de TD&E em conformidade com pressupostos de tecnologia instrucional para que seu novo papel estratégico seja sustentável. Quando considerado o *benchmark* para “E”, o DEA tende a pará-la com ela mesma, haja vista que ela possui um porte muito pequeno frente às demais.

4.4.6 Organização F

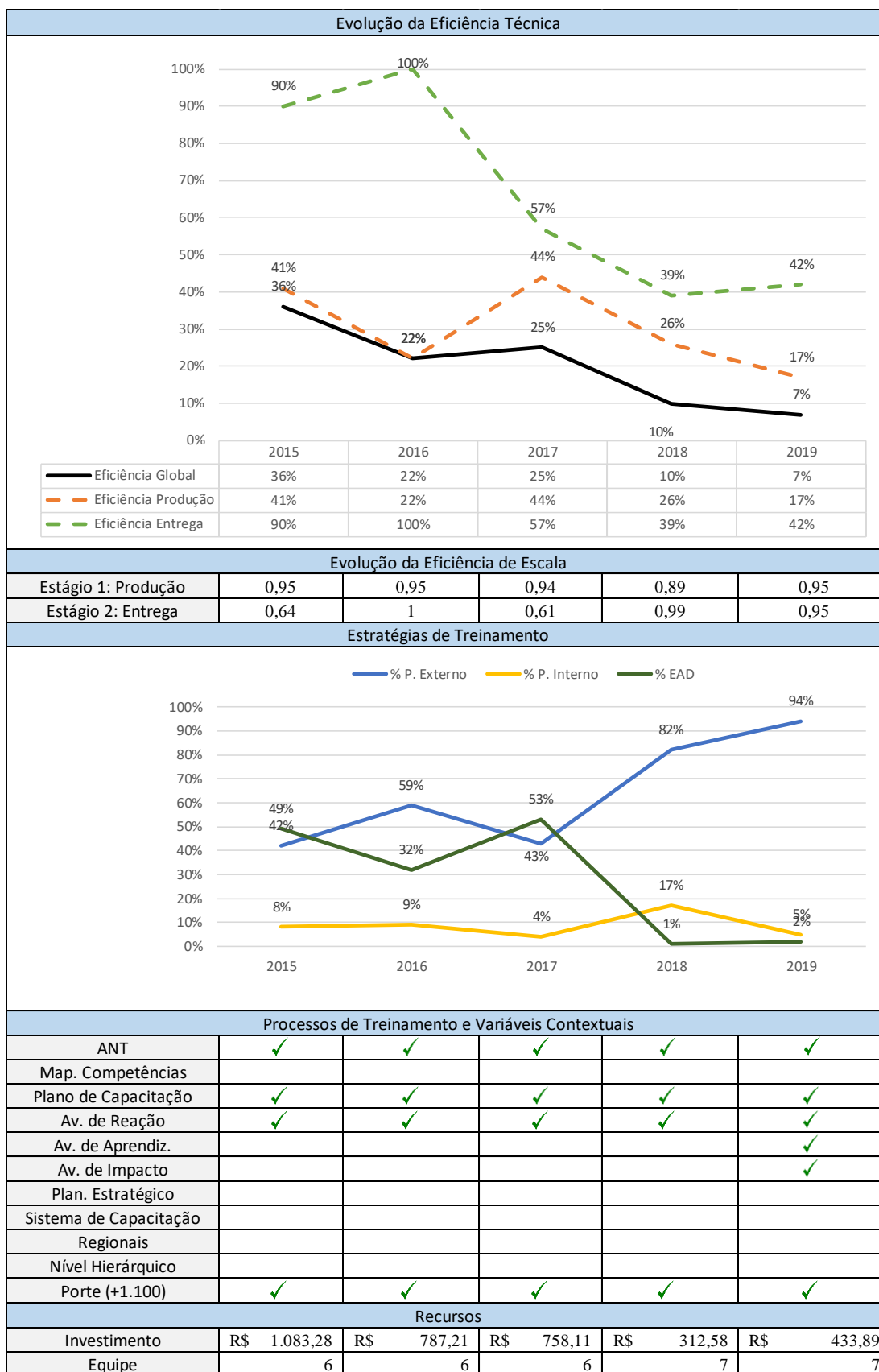


Figura 50. Eficiência e Variáveis contextuais da Organização F

Fonte: Elaboração própria.

A última organização “F” caracteriza-se como uma organização de grande porte da administração direta federal. Ao contrário da unidade de treinamento de “E” que se tornou mais estratégica ao longo do tempo, “F” representa uma tendência oposta que se reflete no declínio de seus níveis de eficiência principalmente a partir de F8 período em que ocorre a mudança de governo que se reflete em mudanças em toda a organização.

A análise dos dados demonstra que a partir desse período houve redução de recursos financeiros e mudança no padrão das estratégias instrucionais que vão se tornando menos diversificadas com o passar do tempo. Com base nos dados de treinamento, em F9 há um padrão similar às unidades de menor desempenho, qual seja a ênfase na contratação de cursos de baixo custo para um menor número de servidores em temas pouco estratégicos.

A despeito da queda de desempenho, o processo de planejamento de “F” consiste em uma metodologia estruturada e aplicada de maneira linear em todo o período. Diferente das demais, a avaliação das necessidades de treinamento é realizada de forma bianual por meio de um processo qualitativo e participativo, em que são realizadas rodadas de reuniões e encontros com chefias, servidores e responsáveis técnicos para que o planejamento da capacitação seja realizado pelas unidades organizacionais. O produto desse processo coletivo termina por ser uma lista de cursos, os quais são priorizados pelas áreas técnicas e sistematizados em um plano pela unidade de treinamento que buscará executá-los. O que se observa em “F” é que essa cultura participativa também está presente na realização de treinamentos em que são comuns ações para grandes números de servidores com a apresentação de palestras e debates. Esses tipos de ações instrucionais são responsáveis pela alta eficiência de entrega em F5 e F6. Ao lado desses cursos, a unidade mantinha uma plataforma corporativa de EAD. Esse cenário muda a partir de F7 com a redução das ações coletivas e mudança no fornecimento do EAD.

Nesse sentido, a queda de eficiência de “F” e as possíveis soluções perpassam pela análise das decisões gerenciais que foram tomadas, e pelas possíveis mudanças que ocorreram no contexto de rotatividade gerencial na organização. Quando analisadas as eficiências de escala, verifica-se que “F” possui poucos desperdícios, atuando em uma escala ideal entre recursos utilizados, treinamentos fornecidos e pessoas capacitadas. Por outro lado, isso é um indicativo de que a ineficiência está mais relacionada à problemas técnicos de execução ou de decisão gerencial. Dessa forma, quando se analisa os *benchmarks* de “F” ela serve de referência para si mesma, devendo, portanto, buscar rever suas estratégias para voltar a atuar como antes.

De forma geral, os achados reportados neste capítulo possibilitaram a análise das atividades de TD&E nas organizações sob uma perspectiva distinta e pouco adotada na literatura, tendo como objetivo a avaliação da eficiência de unidades de treinamento. As avaliações, conforme Borges-Andrade et al. (2012) podem abranger ações educacionais específicas ou programas de treinamento. Conforme verificou-se na revisão de literatura, grande parte dos estudos de avaliação em TD&E têm como enfoque a determinação da efetividade de ações educacionais específicas considerando variáveis relativas ao processo instrucional e a geração de resultados de aprendizagem no nível dos indivíduos e da organização.

A abordagem utilizada nesse estudo, por outro lado, delineou-se para a avaliação dos programas de treinamento, agregando as ações educacionais. Adicionalmente, tendo em vista o desenho comparativo da Análise Envoltória de Dados, a avaliação possibilitou a avaliação de múltiplos programas, de forma conjunta promovendo uma análise da atividade de TD&E tanto em nível organizacional quanto setorial.

Com isso, questões distintas daquelas tratadas por avaliações de ações singulares de treinamento ou avaliações focadas no processo instrucional emergem. A comparação interorganizacional permite verificar, por exemplo, como os programas de treinamento são determinados pelos sistemas organizacionais e institucionais circundantes, possibilitando a compreensão de fatores contextuais que apoiam ou restringem o desempenho das unidades de treinamento. Sobre o aspecto gerencial, a análise comparativa permite avaliar o papel do gestor nesse desempenho ao averiguar os resultados positivos ou negativos de decisões de alocação de recursos e de combinação de alternativas instrucionais.

Dessa forma, a aplicação do DEA contribui com uma perspectiva gerencial do TD&E em que para além das decisões relativas à tecnologia instrucional, são avaliadas as estratégias para a gestão e implementação de programas de treinamento pelas organizações e, para o caso do setor público, pelos gestores de políticas de capacitação no governo. O escopo do modelo operacionalizado nessa pesquisa, demonstrou, portanto, a viabilidade da aplicação da Análise Envoltória de Dados para a investigação do campo de TD&E, trazendo perspectivas de análise complementares aos modelos existentes na literatura.

5 Considerações finais

O presente estudo teve como objetivo principal a proposição de um método para a avaliação da eficiência de unidades de treinamento. Para tanto, buscou a integração dos pressupostos da literatura de avaliação de TD&E com os conceitos oriundos das teorias econômicas de produtividade de forma a estabelecer um conceito de eficiência em treinamento capaz de considerar a complexidade de seus processos e fatores contextuais. Dessa forma, a eficiência é concebida como uma propriedade latente, multidimensional, multivariada e multideterminada do sistema de TD&E. Sob essa perspectiva de eficiência, os métodos baseados na Análise Envoltória de Dados (DEA) se mostram mais adequados ao permitir a consideração de múltiplos insumos, produtos e variáveis contextuais.

A partir de uma revisão sistemática de literatura foram definidas e testadas as variáveis selecionadas para a constituição de um modelo DEA em rede de avaliação da eficiência em dois estágios com abordagem intertemporal. A definição das variáveis guiou a coleta de dados junto a seis organizações públicas federais, compreendendo dados acerca de processos e ações de treinamento realizados no período de 2015 a 2019. A análise possibilitou a classificação das organizações segundo sua capacidade de fornecimento da maior quantidade de horas-aula e oportunidades de treinamento ao menor custo possível. Os modelos e técnicas utilizados na análise produziram índices de eficiência que refletiram o desempenho das unidades, tendo em vista os critérios de economicidade e participação de maior proporção de servidores em ações de treinamento na organização. Adicionalmente, foram empregadas técnicas e análises complementares para o aprofundamento da compreensão acerca dos fatores técnicos, setoriais e organizacionais associados às melhores práticas e às situações de ineficiência observadas.

Conforme demonstrado nesse estudo, o DEA tem ampla aplicação para a avaliação de programas em diversos campos de conhecimento, porém sua utilização para avaliação do Treinamento ainda é escassa e realizada por campos como o da Pesquisa Operacional, Educação e Economia, sem uma conexão clara com a literatura de TD&E. Uma das razões para a escassez de estudos de eficiência em treinamento pode se dever à ausência de bases de dados consolidadas e padronizadas sobre essa atividade, que são mais prevalentes em outras áreas de conhecimento e utilizadas na maioria dos estudos de eficiência. Por outro lado, no âmbito do setor público, recentes alterações da Política Nacional de Desenvolvimento de Pessoas têm visado à sistematização dos dados de treinamento nas organizações públicas federais, de forma que futuros estudos nessa temática poderão ter maior acesso a dados em larga escala, favorecendo a utilização dessa metodologia de avaliação.

Comparativamente à literatura de modelos de avaliação de TD&E, o presente estudo abrangeu um número restrito de variáveis. Logo, ainda que o estudo tenha apresentado uma contribuição metodológica para a avaliação da eficiência de treinamentos, maiores contribuições teóricas oriundas da aplicação do DEA devem necessariamente buscar a inclusão de variáveis tradicionais de eficácia e efetividade do treinamento, relativas aos resultados para indivíduos e organizações. Para tanto, ensejam-se recomendações ao nível da política de capacitação no setor público federal no sentido de se desenvolverem instrumentos para avaliação de resultados do treinamento de forma padronizada entre as organizações, permitindo a sua comparação.

Mesmo com um número relativamente pequeno de variáveis, conclui-se que estas se mostraram adequadas para a produção de índices de eficiência que refletissem dimensões de desempenho relativas aos critérios da economicidade e alcance das ações educacionais implementadas por unidades de treinamento. A comparação sistemática dos casos que compuseram a amostra, sob diferentes perspectivas, modelos e técnicas de análise permitiram o delineamento de melhores práticas, assim como gerar evidências acerca de fatores técnicos, instrucionais, setoriais e organizacionais que afetam o desempenho e a eficiência de unidades de treinamento no setor público.

Com isso, os resultados encontrados demonstraram como as abordagens econômicas para avaliação de eficiência podem ser integradas com os pressupostos teórico-metodológicas do campo de TD&E, assim como fornecer contribuições aos sistemas organizacionais e às políticas de desenvolvimento de pessoas no setor público. O DEA é amplamente utilizado na avaliação de políticas públicas educacionais, econômicas e de saúde (Sant'Anna et al., 2020; Emrouznejad & Yang, 2018), porém possui, conforme revisão de literatura realizada, poucas aplicações no âmbito do treinamento organizacional. Sob o ponto de vista metodológico e procedimental, a pesquisa realiza a aplicação integrada de diferentes modelos de análise da eficiência e documenta os respectivos procedimentos e protocolos de análise realizados em *softwares* estatísticos de acesso livre (Boueri et al., 2015), tornando a metodologia verificável e adaptável a diferentes contextos de pesquisa.

Com relação ao campo de TD&E em específico, verificou-se que a utilização do DEA permite a operacionalização de pressupostos teórico-metodológicos do campo, sobretudo com relação à dinâmica e ao funcionamento dos processos – ou sistemas – de treinamento, cujas prescrições são consolidadas na literatura, mas pouco investigadas empiricamente com relação a seus aspectos sistêmicos, gerenciais e operacionais. Dessa forma, o estudo contribuiu na adoção de uma perspectiva gerencial do treinamento, enfocando a análise em níveis e aspectos

pouco explorados da literatura, tais como avaliação de programas de treinamento em perspectiva comparativa abrangendo múltiplas organizações e a consideração de critérios de avaliação sobre as estratégias de gestão e implementação de programas de treinamento nas organizações, em complemento às perspectivas educacionais e instrucionais mais difundidas. O presente trabalho, portanto, estabelece uma integração entre os campos da Pesquisa Operacional, Economia, TD&E e Gestão Pública de forma a promover contribuições a cada um deles pela utilização de perspectivas teóricas, metodológicas e de escopo de análise que são, em geral, pouco exploradas.

Considera-se, portanto, que a pesquisa alcançou seu principal objetivo, tendo apresentado um método para avaliação da eficiência de unidades de treinamento no setor público federal. A definição de um índice de eficiência global consistiu em uma das principais contribuições metodológicas almejadas pelo estudo, que permite subsidiar avaliações amplas em nível setorial e gerar subsídios para a melhoria da gestão dos treinamentos nas organizações, para monitoramento da implementação e do alcance dos objetivos da Política Nacional de Desenvolvimento de Pessoas em suas dimensões de melhoria da produtividade, economicidade e provisão equânime de oportunidades de treinamento para os servidores públicos federais.

Do ponto de vista profissional, o estudo contribuiu ainda para o avanço metodológico nas áreas de gestão de pessoas e de TD&E no setor público, sob uma perspectiva de gestão baseada em dados e informações. A análise desenvolvida a partir do DEA aplicado às unidades de treinamento no setor público suscita avanços para a otimização, por exemplo, das escalas e portes das unidades de treinamento, possibilitando avaliar a adequação dos recursos humanos e financeiros alocados tendo em vista os objetivos e o contexto de atuação de uma unidade. Assim, torna-se possível avaliar a influência de condicionantes contextuais relevantes para a melhoria do desempenho do treinamento. Por fim, o método proporciona a sistematização de dados acerca das atividades de treinamento que podem subsidiar diretamente decisões por parte de formuladores e implementadores de políticas públicas para gestão e investimento em programas de treinamento em organizações públicas.

Como limitação, pode-se elencar a ênfase na compreensão da eficiência sob uma perspectiva quantitativa, baseada na mensuração das variáveis constituintes do fenômeno (Collis & Hussey, 2005). Consequentemente, deve-se ter em consideração que a avaliação da eficiência realizada abrangeu principalmente os componentes operacionais do processo de treinamento, sem a intenção de se aprofundar em componentes e resultados de caráter qualitativo e que constituem parte do contexto de pesquisa, tais como o valor estratégico das ações educacionais, as percepções, opiniões e perfis de servidores, gestores e dos integrantes

das equipes de treinamento. A consideração de tais aspectos por desenhos mais abrangentes de pesquisa pode ensejar leituras e interpretações distintas acerca dos resultados apresentados neste estudo.

Além dessa questão paradigmática, algumas limitações de ordem metodológica podem ser elencadas, tais como a adstrição do modelo de análise a uma seleção de variáveis de treinamento e a acessibilidade ao campo. O tamanho relativamente reduzido da amostra, ainda que de acordo com o exigido pela literatura e submetido a tratamentos visando minimizar as possíveis distorções analíticas, ensejou a concepção de um modelo de sistema fechado de treinamento, mesmo que o modelo de *Network-DEA* admita a modelagem de um sistema aberto de produção (Cook et al., 2010). Uma perspectiva aberta, com a inclusão de variáveis contextuais seria relevante para os pressupostos teóricos do campo e TD&E que se fundamenta em uma perspectiva sistêmica de influência recíproca com o contexto organizacional (Borges-Andrade & Abbad, 1996).

A análise dos fatores ambientais realizada neste estudo visou abordar as interrelações entre o sistema de treinamento e o contexto organizacional, porém, considera-se que a operacionalização de atividades de treinamento como variáveis categóricas dicotômicas pode ser considerada uma limitação de pesquisa, posto que limita a capacidade de compreender como exatamente essas práticas ocorrem nas organizações, uma vez que são aplicadas de variadas maneiras. Tal questão, portanto, limita a compreensão de como os fatores afetam os resultados do treinamento e sobretudo, como podem se constituir em referencial de melhores práticas para outras organizações. Por fim, outra limitação de natureza metodológica refere-se à utilização da abordagem intertemporal. Verificou-se que tecnologias de diferentes períodos são comparadas diretamente, levando à análise de determinadas firmas com base em tecnologias futuras, o que representa uma limitação à interpretação dos resultados porque determinadas firmas eficientes no passado podem ter sido consideradas ineficientes por tecnologias que surgiriam apenas no futuro.

Tendo em vista as contribuições e as limitações apresentadas pela presente pesquisa, sugere-se a replicação futura do modelo de análise de eficiência proposto em uma amostra de unidades de treinamento que permita a adoção de métodos mais avançados de análise. Nesse ínterim, uma amostra adequada permitirá a adoção de modelos de análise longitudinal mais robustos, tais como os modelos de análises em janela e análise do índice de Malmquist (Ferreira & Gomes, 2020), duas abordagens que permitem a comparação de períodos ao invés da comparação direta de DMUs, detectando melhorias agregadas nas tecnologias de produção.

Com relação à análise de fatores ambientais, sugere-se a exploração de abordagens distintas e complementares à utilizada nesse estudo para determinados tipos de fatores e variáveis não-discrecionárias. Assim, recomenda-se a análise com utilização de regressões estatísticas e a busca de formas de mensuração e avaliação de atividades de treinamento que permitam avaliar a sua influência em índices de eficiência, bem como a adoção de abordagem qualitativa de maneira complementar e mais aprofundada, a fim de melhor compreender e explicar os resultados.

Por fim, cabe ressaltar que variados aspectos do contexto das unidades de treinamento podem ser incorporados à análise para complementação das conclusões obtidas nesse estudo. A avaliação de eficiência pode ser ampliada para que abranja outros tipos de programas de educação formal e desenvolvimento de longo prazo de trabalhadores, assim como operacionalizar medidas que reflitam diferentes estratégias instrucionais, fatores ambientais, medidas qualitativas de valor e, principalmente, a extensão do modelo pela adição de estágios subsequentes do processo de treinamento referentes aos estágios de resultados de aprendizagem, reação e impacto ou mesmo resultados organizacionais e sociais em diferentes níveis e perspectivas (Aguinis & Kraiger, 2009), os quais podem ser integrados como medidas de valor final a um modelo de *Network-DEA* multiestágio.

Referências

- Abbad, G. (1999) Um Modelo Integrado de Avaliação de Impacto do Treinamento no Trabalho: IMPACT. 262 f. Tese (Doutorado em Psicologia) – Instituto de Psicologia, Universidade de Brasília, Brasília.
- Abbad, G., Gama, A. L. G., & Borges-Andrade, J. E. (2000). Treinamento: análise do relacionamento da avaliação nos níveis de reação, aprendizagem e impacto no trabalho. *Revista de Administração Contemporânea*, 4(3), 25-45. <https://doi.org/10.1590/S1415-65552000000300003>
- Abbad, G. S.; Borges-Andrade, J. E. (2014) *Aprendizagem humana em organizações de trabalho*. In: Zanelli, J. C.; Borges-Andrade, J. E.; Bastos, A. V. B. (Eds.). *Psicologia, organizações e trabalho no Brasil (2ª Ed.)*. Porto Alegre: Artmed, p. 237-275, 2014.
- Abbad, G. S., Nascimento, A. S., & Gaspar, F. D. R. (2020). Avaliação em treinamento, desenvolvimento e educação corporativa. In: Hutz, C. S., Bandeira, D. R., Trentini, C. M., & Vazquez, A. C. S. (Eds.). *Avaliação psicológica no contexto organizacional e do trabalho (pp. 226-246)*. Porto Alegre: Artmed.
- Abbad, G., Pilati, R., & Pantoja, M. J. (2003). Avaliação de treinamento: análise da literatura e agenda de pesquisa. *RAUSP Management Journal*, 38(3), 205-218.
- Afonso, A. (2007). A Eficiência do Estado. In R. Boueri & M. Saboya (Eds.), *Aspectos do Desenvolvimento Fiscal* (pp. 111–124). Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA).
- Aguinis, H., & Kraiger, K. (2009). Benefits of training and development for individuals and teams, organizations, and society. *Annual Review of Psychology*, 60, 451–474. <https://doi.org/10.1146/annurev.psych.60.110707.163505>
- Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shotland, A. (1997). Meta-Analysis Relations among Training Criteria. *Personnel Psychology*, 50(2), 341–358. <https://doi.org/10.1111/j.1744-6570.1997.tb00911.x>
- Alliger, G. M., & Janak, E. A. (1989). *Kirkpatrick 'S Levels Of Training Criteria : Thirty Years Later*.
- Alvarez, K., Salas, E., & Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. *Human resource development Review*, 3(4), 385-416.
- Aviles-Sacoto, S., Cook, W. D., Imanirad, R., & Zhu, J. (2015). Two-stage network DEA: when intermediate measures can be treated as outputs from the second stage. *Journal of the Operational Research Society*, 66(11), 1868-1877.
- Bach, S. (2010). Human resource management in public sector. In: Wilkinson, Adrian et al. *The sage handbook of human resource management*. Londres: Sage Publications.

- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30(9), 1078–1092. <https://doi.org/10.1287/mnsc.30.9.1078>
- Baraldi, S., & Cifalinò, A. (2015). Delivering training strategies: The balanced scorecard at work. *International Journal of Training and Development*, 19(3), 179–198. <https://doi.org/10.1111/ijtd.12055>
- Barbosa, F. C., & Fuchigami, H. Y. (2018). *Análise Envoltória de Dados: Teoria e Aplicações* (1st ed.). Itumbiara: GO: ULBRA.
- Becker, B. E., Huselid, M. A., & Ulrich, D. (2001). *Gestão estratégica de pessoas com Scorecard: interligando pessoas, estratégia e performance* (8th ed.). Rio de Janeiro: Elsevier.
- Bell, B. S., Tannenbaum, S. I., Ford, J. K., Noe, R. A., & Kraiger, K. (2017). 100 years of training and development research: What we know and where we should go. *Journal of Applied Psychology*, 102(3), 305–323. <https://doi.org/10.1037/apl0000142>
- Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R* (Vol. 157). Springer New York. <https://doi.org/10.1007/978-1-4419-7961-2>
- Bogetoft, P., & Otto, L. (2019) Benchmark and frontier analysis using DEA and SFA. [s.l.] : R. package.
- Bogetoft, P., & Wittrup, J. (2017). Apples to oranges: benchmarking vocational education and training programmes. *School Effectiveness And School Improvement*, 28(2), 191–216. <https://doi.org/10.1080/09243453.2016.1247731>
- Borges-Andrade, J. E. (2006a). Avaliação integrada e somativa em TD&E. In J. E. Borges-Andrade, G. D. S. Abbad, & L. Mourão (Eds.), *Treinamento, Desenvolvimento e Educação em Organizações e Trabalho: Fundamentos para a Gestão de Pessoas*. (pp. 343–358). Porto Alegre: Artmed.
- Borges-Andrade, J. E. (2006b). Competência Técnica e Política do Profissional de TD&E. In J. E. Borges-Andrade, G. D. S. Abbad, & L. Mourão (Eds.), *Treinamento, Desenvolvimento e Educação em Organizações e Trabalho: Fundamentos para a Gestão de Pessoas*. (pp. 177–195). Porto Alegre: Artmed.
- Borges-Andrade, J. E. (2002). Desenvolvimento de medidas em avaliação de treinamento TT - Development of measures in training evaluation. *Estudos de Psicologia (Natal)*, 7(spe), 31–43. <https://doi.org/10.1590/S1413-294X2002000300005>
- Borges-Andrade, J. E., & Abbad, G. da S. (1996). Treinamento e desenvolvimento: reflexões sobre suas pesquisas científicas. *Revista de Administração*, 31(2), 112–125.
- Borges-Andrade, J. E., Abbad, G. da S., & Mourão, L. (2012). Modelos de avaliação e aplicação em TD&E. In G. da S. Abbad, L. Mourão, P. P. M. Meneses, T. Zerbini, J. E. B. -Andrade, & Raquel Vilas-Boas (Eds.), *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação* (pp. 20–35). São Paulo: Artmed.

- Borges-Andrade, J. E., & Pagotto, C. do P. (2010). O estado da arte da pesquisa brasileira em Psicologia do Trabalho e Organizacional. *Psicologia: Teoria e Pesquisa*, 26(spe), 37–50. <https://doi.org/10.1590/S0102-37722010000500004>
- Borges-Andrade, J. E., Zerbini, T., Abbad, G. da S., & Mourão, L. (2013). Treinamento, Desenvolvimento e Educação: Um Modelo para sua Gestão. In L. de O. Borges & L. Mourão (Eds.), *O Trabalho e as Organizações: Atuações a partir da Psicologia* (pp. 466–496). Porto Alegre: Artmed.
- Borges-Andrade, J. E. (1982). Avaliação somativa de sistemas institucionais: integração de três propostas. *Tecnologia Educacional*, Rio de Janeiro, v.11, n. 46.
- Boueri, R., Rocha, F., & Rodopoulos, F. (Eds.). (2015). Avaliação da qualidade do gasto público e mensuração da eficiência. Ministério da Fazenda, Secretaria do Tesouro Nacional.
- Brandão, H. P. (2007). Competências no trabalho: uma análise da produção científica brasileira. *Estudos de Psicologia (Natal)*, 12(2), 149–158. <https://doi.org/10.1590/S1413-294X2007000200007>
- Brasil. (2019a). *Decreto Nº 9.991, De 28 De Agosto De 2019 - Dispõe sobre a Política Nacional de Desenvolvimento de Pessoas da administração pública federal direta, autárquica e fundacional, e regulamenta dispositivos da Lei nº 8.112, de 11 de dezembro de 1990, quanto a l.* Brasília: Presidência da República.
- Brasil. (2019b). *Instrução Normativa Nº 201, De 11 De Setembro De 2019 Dispõe sobre os critérios e procedimentos específicos para a implementação da Política Nacional de Desenvolvimento de Pessoas, de que trata o Decreto nº 9.991, de 28 de agosto de 2019, pelos órgãos int.* Brasília: Ministério da Economia.
- Brasil. (2006). *Decreto Nº 5.707, De 23 De Fevereiro De 2006 - Institui a Política e as Diretrizes para o Desenvolvimento de Pessoal da administração pública federal direta, autárquica e fundacional, e regulamenta dispositivos da Lei nº 8.112, de 11 de dezembro de 1990.* Brasília: Presidência da República
- Brauchle, P. E., & Schmidt, K. (2007). Contemporary Approaches for Assessing Outcomes on Training, Education, and HRD Programs. *Journal of Industrial Teacher Education*, 41(3), 1–17.
- Bretschneider, S., Marc-Aurele, F. J., & Wu, J. (2004). “Best practices” research: a methodological guide for the perplexed. *Journal of Public Administration Research and Theory*, 15(2), 307-323.
- Brinkerhoff, R. O. (1998). Clarifying and directing impact evaluation. In *Evaluating corporate training: Models and issues* (pp. 141-166). Springer, Dordrecht.
- Brinkerhoff, R. O. (2003). *The success case method: Find out quickly what's working and what's not.* Berrett-Koehler Publishers.

- Buleca, J., & Mura, L. (2014). Quantification of the Efficiency of Public Administration by Data Envelopment Analysis. *Procedia Economics and Finance*, 15(14), 162–168. [https://doi.org/10.1016/s2212-5671\(14\)00469-9](https://doi.org/10.1016/s2212-5671(14)00469-9)
- Bushnell, D. S. (1990). Input, process, output: A model for evaluating training. *Training & Development Journal*, 44(3), 41-44.
- Camões, M. R. de S., & Meneses, P. P. M. (2016). *Gestão de Pessoas no Governo Federal: Análise da Implementação da Política Nacional de Desenvolvimento de Pessoal* (Cadernos,). Brasília: ENAP.
- Campbell, J. P. (1971). Personnel training and development. *Annual review of psychology*, 22(1), 565-602.
- Casado, F. L. (2007). Análise Envoltória De Dados: Conceitos, Metodologia E Estudo Da Arte Na Educação Superior. *Revista Sociais e Humanas*, 20(1), 59–71. <https://doi.org/10.5902/23171758907>
- Cassiano, S. K., & Borges-Andrade, J. E. (2017). Methodological delimitations of the Brazilian research on effects of training. *Revista Psicologia Organizações e Trabalho*, 17(3), 141-149.
- Cavalcante, P., & Carvalho, P. (2017). Profissionalização da burocracia federal brasileira (1995-2014): avanços e dilemas. *Revista de Administração Pública*, 51(1), 1–26. <https://doi.org/10.1590/0034-7612144002>
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- Charnes, A., Cooper, W. W., & Rhodes, E. (1981). Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through. *Management science*, 27(6), 668-697.
- Chartered Institute of Personnel And Development - CIPD. (2019) Addressing Employer Underinvestment in Training: the case for a broader training lev,; Report 2019 [online]. London: CIPD.
- Chyung, S. Y. (2019). *10-Step Evaluation for Training and Performance Improvement*. Singapore: Sage Publications.
- Collis, J., & Hussey, R. (2005). *Pesquisa em administração: um guia prático para alunos de graduação e pós-graduação*. Bookman.
- Cook, W. D., Zhu, J., Bi, G., & Yang, F. (2010). Network DEA: Additive efficiency decomposition. *European Journal of Operational Research*, 207(2), 1122–1129. <https://doi.org/10.1016/j.ejor.2010.05.006>

- Cooper, W. W., Seiford, L. M., & Tone, K. (2007). *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA Solver Software. Data Envelopment Analysis.*
- Cooper, W. W., Seiford, L. M., & Zhu, J. (2011). Handbook on Data Envelopment Analysis. *Chapter 1: Data Envelopment Analysis.*
- Cordero, J. M., Santín, D., & Simancas, R. (2017). Assessing European primary school performance through a conditional nonparametric model. *Journal of the Operational Research Society*, 68(4), 364-376.
- Côrtes, F. G., & Meneses, P. P. M. (2019). Gestão estratégica de pessoas no Legislativo Federal brasileiro: condições para a implementação. *Revista de Administração Pública*, 53(4), 657-686.
- Costa, E. M., de Souza, H. R., Ramos, F. de S., & da Silva, J. L. M. (2012). Eficiência e desempenho no ensino superior: Uma análise da fronteira de produção educacional das ifes brasileiras. *Revista de Economia Contemporanea*, 16(3), 415-440. <https://doi.org/10.1590/S1415-98482012000300003>
- Chochard, Y., & Davoine, E. (2011). Variables influencing the return on investment in management training programs: a utility analysis of 10 Swiss cases. *International Journal of Training and Development*, 15(3), 225-243. <https://doi.org/10.1111/j.1468-2419.2011.00379.x>
- Cubbin, J., & Zamani, H. (1996). A comparison of performance indicators for training and enterprise councils in the UK. *Annals of Public and Cooperative Economics*, 67(4), 603-632. <https://doi.org/10.1111/j.1467-8292.1996.tb01410.x>
- Damasceno, S. M. S., Abbad, G., & Meneses, P. P. M. (2012). Modelos lógicos e avaliações de treinamentos organizacionais. *Paidéia (Ribeirão Preto)*, 22, 217-228.
- Debnath, R. M., Shankar, R., & Kumar, S. (2008). Performance evaluation of technical institutions: an application of data envelopment analysis. *Education, Knowledge & Economy*, 2(1), 51-66.
- Debnath, R. M., & Shankar, R. (2009). Assessing performance of management institutions: An application of data envelopment analysis. *The TQM Journal*.
- Dessinger, J. C., & Moseley, J. L. (2010) Full-Scope Evaluation: Do You "Really Oughta, Wanna"?. In: Moseley, J. L., & Dessinger, J. C. (Eds). Handbook of improving performance in the workplace. Volume 3, Measure and Evaluation. Wiley/Pfeiffer.
- De Witte, K., & López-Torres, L. (2017). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*, 68(4), 339-363. <https://doi.org/10.1057/jors.2015.92>
- Dutta, P., Jaikumar, B., & Arora, M. S. (2021). Applications of data envelopment analysis in supplier selection between 2000 and 2020: a literature review. *Annals of Operations Research*, 1-56.

- Ekiz, M. K., & Tuncer Sakar, C. (2020). A new DEA approach to fully rank DMUs with an application to MBA programs. *INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH*, 27(4, SI), 1886–1910. <https://doi.org/10.1111/itor.12635>
- Emrouznejad, A., Banker, R., Miranda Lopes, A. L., & Rodrigues de Almeida, M. (2014). Data Envelopment Analysis in the public sector. *Socio-Economic Planning Sciences*, 48(1), 2–3. <https://doi.org/10.1016/j.seps.2013.12.005>
- Emrouznejad, A., & Yang, G. liang. (2018). A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016. *Socio-Economic Planning Sciences*, 61, 4–8. <https://doi.org/10.1016/j.seps.2017.01.008>
- ENAP. (2020). *Competências transversais de um setor público de alto desempenho* (p. 28). p. 28. Brasília: Escola Nacional de Administração Pública (Enap).
- Falquetto, A. M., Takasago, M., Peña, C. R., de Araújo Neto, L. M., & Sales, I. C. H. (2018). Avaliação da eficiência dos programas de economia no país contemplados com o Proex e o Proap. *Race: revista de administração, contabilidade e economia*, 17(1), 333-364.
- Fávero, L. P. & Belfiore, P. B. (2012). *Pesquisa operacional: para cursos de administração, contabilidade e economia*. Elsevier.
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES A-STATISTICS IN SOCIETY*, 120(3), 294.
- Ferguson, B. (2019). Public administration performance measurement from Competing for Influence: The Role of the Public Service in Better Government in Australia on JSTOR. In *Competing for Influence: The Role of the Public Service in Better Government in Australia* (pp. 127–166). Retrieved from https://www.jstor.org/stable/j.ctvp2n3pr.9?seq=1#metadata_info_tab_contents
- Ferreira, C. M. de C., & Gomes, A. P. (2020). *Introdução à Análise Envoltória de Dados: Teorias, Modelos e Aplicações*. (2nd ed.). Viçosa: MG: UFV.
- Ferreira, & Abbad, G. da S. (2014). Avaliação de necessidades de treinamento no trabalho: ensaio de um método prospectivo. *Revista Psicologia*, 14(1), 01–17. Retrieved from http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1984-66572014000100002&lng=pt&nrm=iso&tlng=pt
- Firmino, M. J. A. C. S. (2015). *Teste de hipóteses: uma abordagem não paramétrica*. (Dissertação de mestrado). Universidade de Lisboa, Portugal.
- Fisher, D. M., Kiang, M., & Fisher, S. A. (2007). A value-added approach to selecting the best master of business administration (MBA) program. *Journal of Education for Business*, 83(2), 72-76.
- Fisher, S., Chi, R., Fisher, D., & Kiang, M. (2017). Determining the value of undergraduate business programs from market vs academic perspectives. *International Journal of Educational Management*.

- Flick, U. (2009). *Qualidade na pesquisa qualitativa: coleção pesquisa qualitativa*. Bookman Editora.
- Fonseca, D. R., Meneses, P. P. M., Silva, A. I., Filho, & Campos, N. G. (2013). Autonomia para gestão estratégica de pessoas no setor público federal: perspectivas de análise e agenda de pesquisa. *Revista de Administração Pública*, 47(6), 1451-1475.
- Fonseca, D. R., & Meneses, P. P. M. (2016). Fatores para implantação e desenvolvimento da gestão por competências em agências reguladoras federais. *Revista Eletrônica Científica da UERGS*, 2(2), 117-133.
- Fonseca, D. R. da, & Meneses, P. P. M. (2017). Transição Burocrática-Estratégica de Atividades de Treinamento em Organizações Públicas Federais: Uma Análise com Redes Sociais de Afiliação. *XLI Encontro Da ANPAD*, 1–16.
- Fonseca, D. R. da, & Meneses, P. P. M. (2018). Gestão de Redes de Capacitação: metodologia para um modelo de governança. *Cadernos ENAP*, 54, 187.
- Ford, D. J. (1993). Benchmarking HRD. *Training & Development*, 47(6), 36-42.
- Forsund, F. R., & Sarafoglou, N. (2002). On the origins of data envelopment analysis. *Journal of Productivity Analysis*, 17(1), 23-40.
- Freitas, I. A., & Borges-andrade, J. E. (2004). Efeitos de treinamento nos desempenhos individual e organizacional. *Revista de Administração de Empresas*, 44(3), 44–56. Retrieved from <http://www.spell.org.br/documentos/ver/11079/efeitos-de-treinamento-nos-desempenhos-individual-e-organizacional>
- Freitas, I. A., & Mourão, L. (2012). Medidas de impacto de TD&E: o nível da mudança organizacional. In G. da S. Abbad, L. Mourão, P. P. M. Meneses, T. Zerbini, J. E. B. -Andrade, & Raquel Vilas-Boas (Eds.), *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação* (pp. 163–176). São Paulo: Artmed.
- Freitas, W. R. de S., Jabbour, J. C. C., & Santos, F. C. A. (2011). Continuing the evolution: towards sustainable HRM and sustainable organizations. *Business Strategy Series*, 12(5), 226–234. <https://doi.org/10.1108/17515631111166861>
- Gagne, R. M. (1970). *The conditions of learning* (2nd ed.). Holt, Rinehart & Winston.
- Gagnon, M.-E., & Smith, M. (2013). The Effects of a Training Levy on Training Characteristics and Outcomes: The Case of Quebec. *Relations Industrielles / Industrial Relations*, 68(1), 120–141. Retrieved from <http://www.jstor.org/stable/23646209>
- Garavan, T. N. (1997) Training, development, education and learning: different or the same? *Journal of European Industrial Training*, 21(2), p. 39-50.
- Garavan, T. N.; Heraty, N.; Barnicle, B. (1999) Human resource development literature: current issues, priorities and dilemmas. *Journal of European Industrial Training*, 23(4), 169-179.

- Garavan, T. N., McCarthy, A., Sheehan, M., Lai, Y., Saunders, M. N. K., Clarke, N., ... Shanahan, V. (2019). Measuring the organizational impact of training: The need for greater methodological rigor. *Human Resource Development Quarterly*, 30(3), 1–19. <https://doi.org/10.1002/hrdq.21345>
- Giacomello, C. P., & de Oliveira, R. L. (2014). Análise Envoltória de Dados (DEA): uma proposta para avaliação de desempenho de unidades acadêmicas de uma universidade. *Revista Gestão Universitária na América Latina-GUAL*, 130-151.
- Golany, B., & Roll, Y. (1989). An application procedure for DEA. *Omega*. [https://doi.org/10.1016/0305-0483\(89\)90029-7](https://doi.org/10.1016/0305-0483(89)90029-7)
- Goldstein, I. L. (1980). Training in Work Organizations. *Annual Review of Psychology*, 31(1), 229–272. <https://doi.org/10.1146/annurev.ps.31.020180.001305>
- Gondim, S. M. G., Borges-Andrade, J. E., & Bastos, A. V. B. (2019). Desenvolvimento científico e desafios da Psicologia do Trabalho e das Organizações no Brasil. *Estudos e Pesquisas Em Psicologia*, 18(4), 1087–1105. <https://doi.org/10.12957/epp.2018.42224>
- Gratton, L. (1999) People process as a source of competitive advantage. In: L. Gratton, l.; Hailey, v. H.; Stiles, p.; Truss, C. *Strategic human resource Management: Corporate rethoric and human reality*. Oxford: University Press, 170-198.
- Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2014a). Efficiency in education: Research and implications. *Applied Economic Perspectives and Policy*, 36(2), 175-210.
- Grosskopf, S., Hayes, K. J., & Taylor, L. L. (2014b). Applied efficiency analysis in education. *Economics and Business Letters*, 3(1), 19-26.
- Haksever, C., & Muragishi, Y. (1998). Measuring value in MBA programmes. *Education Economics*, 6, 11-26.
- Hamblin, A. C. (1978). *Avaliação e controle do treinamento*. São Paulo: McGraw-Hill.
- Hirao, Y. (2012). Efficiency of the top 50 business schools in the United States. *Applied Economics Letters*, 19(1), 73-78.
- Holton, E. F. (1996). The Flawed Four-Level Evaluation Model. *Human Resource Development Quarterly*, 7(1), 5–21. <https://doi.org/10.1002/hrdq.3920070103>
- Hsu, M. K., James, M. L., & Chao, G. H. (2009). An efficiency comparison of MBA programs: top 10 versus non-top 10. *Journal of Education for Business*, 84(5), 269-274.
- Iyer, K. C., & Jain, S. (2019). Performance measurement of airports using data envelopment analysis: A review of methods and findings. *Journal of Air Transport Management*, 81, 101707.
- Jannuzzi, P. D. M. (2005). Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil.

- Johnes, J., Portela, M., & Thanassoulis, E. (2017). Efficiency in education. *Journal of the Operational Research Society*, 68(4), 331–338. <https://doi.org/10.1057/s41274-016-0109-z>
- Kao, C. (2017). *Network Data Envelopment Analysis* (Vol. 240). Springer International Publishing. <https://doi.org/10.1007/978-3-319-31718-2>
- Kao, L.-J., Chiu, S.-Y., & Ko, H.-T. (2014). A Study of the Talent Training Project Management for Semiconductor Industry in Taiwan: The Application of a Hybrid Data Envelopment Analysis Approach. *SCIENTIFIC WORLD JOURNAL*. <https://doi.org/10.1155/2014/296345>
- Kaplan, R. S., & Norton, D. P. (1996). *Translating strategy introduction the balanced scorecard*. Harvard Business School.
- Kast, F. E., & Rosenzweig, J. E. (1972). General Systems Theory: Applications for Organization and Management. *Academy of Management Journal*, 15(4), 447–465.
- Katz, D., & Kahn, R. L. (1978). *Psicologia Social das Organizações* (2nd ed.). São Paulo: Atlas.
- Kaufman, R., & Keller, J. M. (1994). Levels of Evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, 5(4), 371–380. Retrieved from <https://search.proquest.com/docview/61576689?accountid=26646>
- Kennedy, P. E., Chyung, S. Y., Winiecki, D. J., & Brinkerhoff, R. O. (2014). Training professionals' usage and understanding of Kirkpatrick's Level 3 and Level 4 evaluations. *International Journal of Training and Development*, 18(1), 1–21. <https://doi.org/10.1111/ijtd.12023>
- Kellaghan, T., & Stufflebeam, D. L. (Eds.). (2003). *International Handbook of Educational Evaluation: Part One: Perspectives/Part Two: Practice* (Vol. 9). Springer Science & Business Media.
- Kellogg Foundation. (2004). *WK Kellogg Foundation Logic Model Development Guide*. WK Kellogg Foundation.
- Kirkpatrick, D. L. (1976) Evaluation of training. In: Craig, R. L. *Training and Development handbook*. 2 ed. New York: McGraw-Hill, pp. 18.1-18.27.
- Kohl, S., Schoenfelder, J., Fügener, A., & Brunner, J. O. (2019). The use of Data Envelopment Analysis (DEA) in healthcare with a focus on hospitals. *Health care management science*, 22(2), 245-286.
- Kong, W.-H., & Fu, T.-T. (2012). Assessing the performance of business colleges in Taiwan using data envelopment analysis and student based value-added performance indicators. *OMEGA-INTERNATIONAL JOURNAL OF MANAGEMENT SCIENCE*, 40(5), 541–549. <https://doi.org/10.1016/j.omega.2011.10.004>

- Kraiger, K., & Ford, J. K. (2021). The science of workplace instruction: Learning and development applied to work. *Annual Review of Organizational Psychology and Organizational Behavior*, 8, 45-72.
- Kraiger, K., Ford, J. K., & Salas, E. (1993). *Application of Cognitive , Skill-Based , and Affective Theories of Learning Outcomes to New Methods of Training Evaluation*. 78(2).
- Kraiger, K., McLinden, D., & Casper, W. J. (2004). Collaborative planning for training impact. *Human Resource Management*, 43(4), 337–351.
<https://doi.org/10.1002/hrm.20028>
- Kumar, S. (2019). Artificial intelligence divulges effective tactics of top management institutes of India. *Benchmarking- An International Journal*, 26(7), 2188–2204.
<https://doi.org/10.1108/BIJ-08-2018-0251>,
- Kumar, N., & Singh, A. (2014). Efficiency analysis of banks using DEA: A review. *International Journal of Advance Research and Innovation*, 1, 120-126.
- Langbein, L., & Felbinger, C. L. (2006). *Public Program Evaluation: A Statical Guide*. New York: MB Sharpe Inc.
- Latham, G. (1988). Human Resource Training And Development. *Annual Review of Psychology*, 39(1), 545–582. <https://doi.org/10.1146/annurev.psych.39.1.545>
- Lindsay, A. W. (1982). Institutional performance in higher education: The efficiency dimension. *Review of Educational Research*, 52(2), 175-199.
- Loiola, E., N ris, J. S., & Bastos, A. V. B. (2006). Aprendizagem em organiza es: mecanismos que articulam processos individuais e coletivos. In J. E. Borges-Andrade, G. da S. Abbad, & L. Mour o (Eds.), *Treinamento, Desenvolvimento e Educa o em Organiza es e Trabalho: Fundamentos para a Gest o de Pessoas* (pp. 114–136). Porto Alegre: Artmed.
- Lovell, C. A. K. (1993). Production frontiers and productive efficiency. In H. O. Fried, C. A. K. Lovell, & S. S. Schmidt (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications* (p. 426). Oxford: Oxford University Press.
- Markoulli, M. P., Lee, C. I. S. G., Byington, E., & Felps, W. A. (2017). Mapping Human Resource Management: Reviewing the field and charting future directions. *Human Resource Management Review*, 27(3), 367–396.
<https://doi.org/10.1016/j.hrmmr.2016.10.001>
- Manzoor, A. (2014). A look at efficiency in public administration: Past and future. *SAGE Open*.
<https://doi.org/10.1177/2158244014564936>
- Mariano, E. B., Almeida, M. R., & Rebelatto, D. a. N. (2006). Peculiaridades da An lise por Envolt ria de Dados. *Xii Simpep*, (November), 11. <https://doi.org/10.13140/2.1.4168.8000>
- Mariano, E. B. (2007). Conceitos B sicos de An lise de Efici ncia produtiva. *XIV Simp sio de Engenharia Da Produ o*, 1–12. Bauru: SP: XIV SIMPEP.

- McLaughlin J. A., & Jordan, G. B. (2015). Using Logic Models. In: Newcomer, K. E., Hatry, H. P., & Wholey, J. S (Eds.). *Handbook of practical program evaluation*. John Wiley & Sons.
- Mahmoudi, R., Emrouznejad, A., Shetab-Boushehri, S. N., & Hejazi, S. R. (2020). The origins, development and future directions of data envelopment analysis approach in transportation systems. *Socio-Economic Planning Sciences*, 69, 100672.
- Mardani, A., Zavadskas, E. K., Streimikiene, D., Jusoh, A., & Khoshnoudi, M. (2017). A comprehensive review of data envelopment analysis (DEA) approach in energy efficiency. *Renewable and Sustainable Energy Reviews*, 70, 1298-1322.
- Mariano, E. B., Sobreiro, V. A., & do Nascimento Rebelatto, D. A. (2015). Human development and data envelopment analysis: A structured literature review. *Omega*, 54, 33-49.
- Mark, M. M., Henry, G. T., and Julnes, G. (2000) *Evaluation: An Integrated Framework for Understanding, Guiding, and Improving Policies and Programs*. San Francisco: JosseyBass.
- Mauss, C. V., Flach, F., & Bleil, C. (2012). A informação de custos como base das decisões de gestão financeira no setor público. *XIX Congresso Brasileiro de Custos*, 1–15. Bento Gonçalves, RS, Brasil.
- Mehra, A., Langer, N., Bapna, R., & Gopal, R. (2014). Estimating Returns to Training in the Knowledge Economy. *MIS Quarterly*, 38(3), 757–772. <https://doi.org/10.2307/26634992>
- Meneses, P. P. M., & Abbad, G. (2009). Proposta para Desenvolvimento de Modelos de Avaliação da Efetividade de Programas de Treinamento. *RAC-Eletrônica*, 3(1), 105–122. Retrieved from <http://www.spell.org.br/documentos/ver/31130/proposal-for-the-development-of-models-to-evaluate-the-effectiveness-of-training-programs>
- Meneses, Zerbini, T., & Abbad, G. da S. (2010). *Manual De Treinamento Organizacional*. Porto Alegre: Artmed.
- Modesto, P. (2000). Notas para um debate sobre o princípio da eficiência. *Revista Do Serviço Público*, 51(2), 105–119. <https://doi.org/10.21874/rsp.v51i2.328>
- Mourão, L., & Borges-Andrade, J. E. (2005). Avaliação de programas públicos de treinamento : um estudo sobre o impacto no trabalho e na geração de emprego. *Organizações & Sociedade*, 12(33), 13–38. <https://doi.org/10.1590/S1984-92302005000200001>
- Mourão, L. & Meneses (2012). Marco lógico como ferramenta de avaliação em TD&E. In G. da S. Abbad, L. Mourão, P. P. M. Meneses, T. Zerbini, J. E. B. -Andrade, & Raquel Vilas-Boas (Eds.), *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação* (pp. 36–49). São Paulo: Artmed.
- Mourão, L., Porto, J. B., & Puente-Palacios, K. (2014). Construção e evidências de validade de duas escalas de percepção de desenvolvimento profissional. *Psico-USF*, 19(1), 73–85. <https://doi.org/10.1590/S1413-82712014000100008>

- Nadler, L. (1984). Human Resource Development. In NADLER, L; NADLER, Z. (ed.) *Handbook of human resource development*. John Wiley & Sons, 1984.
- Nascimento, A. de S. (2018). Efeitos de treinamentos no desempenho de gestores: o caso das cooperativas de crédito. (Dissertação de mestrado). Universidade de Brasília, Brasil.
- Ng, Y. C., & Li, S. (2003). Training and Productive Efficiency in Transition Economies. *Oxford Economic Papers*, 55(4), 607–624. Retrieved from <http://www.jstor.org/stable/3488789>
- Nojavan, M., Heidari, A., & Mohammaditabar, D. (2020). A fuzzy service quality based approach for performance evaluation of educational units. *Socio-Economic Planning Sciences*, 100816. <https://doi.org/https://doi.org/10.1016/j.seps.2020.100816>
- OCDE. (2010). *Avaliação da Gestão de Recursos Humanos no Governo - Relatório da OCDE: Brasil 2010*.
- OCDE. (2017). *National Schools of Government. Building Civil Service Capacity*. OECD Report: Paris.
- Odelius, C. C., & de Siqueira Júnior, F. A. B. (2011). Construção e validação de instrumento de pesquisa relativo a aspectos que influenciam a efetividade de políticas de TD&E. *Journal of Accounting, Management and Governance*, 14(1).
- Pacheco, R. S. (2002). Human resources policy for the managerial renovation: accomplishments between 1995-2002. *Revista Do Serviço Público*, 53(4), 79–106. Retrieved from <http://www.spell.org.br/documentos/ver/35895/human-resources-policy-for-the-managerial-renovation--accomplishments-between-1995-2002>
- Pagani, R. N., Kovalski, J. L., & Resende, L. M. (2015). Methodi Ordinatio: a proposed methodology to select and rank relevant scientific papers encompassing the impact factor, number of citation, and year of publication. *Scientometrics*, 105(3), 2109–2135. <https://doi.org/10.1007/s11192-015-1744-x>
- Pai, P., Khan, B. M., & Mukherjee, P. N. Data Envelopment Analysis (DEA)–Application at NMIMS–SBM, a Leading AACSB Accredited Indian Higher Education Business School.
- Passmore, J., & Velez, M. J. (2012). *SOAP-M: a training evaluation model for HR*. (August). <https://doi.org/10.1108/00197851211254743>
- Passmore, J., & Velez, M. J. (2015). Chapter 8: Training Evaluation. In K. Kraiger, J. Passmore, N. R. dos Santos, & S. Malvezzi (Eds.), *The Wiley Blackwell Handbook of the Psychology of Training, Development, and Performance Improvement* (1st ed., pp. 136–153). John Wiley & Sons.
- Pelone, F., Kringos, D. S., Romaniello, A., Archibugi, M., Salsiri, C., & Ricciardi, W. (2015). Primary care efficiency measurement using data envelopment analysis: a systematic review. *Journal of medical systems*, 39(1), 1-14.
- Phillips, J. J. (1991). Measuring the return on HRD. *Employment Relations Today*, 18(3), 329-342.

- Phillips, P. P., & Phillips, J. J. (2010) Return on Investment. In: Moseley, J. L., & Dessinger, J. C. (Eds). *Handbook of improving performance in the workplace. Volume 2.* Wiley/Pfeiffer.
- Pilati, R. (2006). História e importância de TD&E. In J. E. Borges-Andrade, G. da S. Abbad, & L. Mourão (Eds.), *Treinamento, Desenvolvimento e Educação em Organizações e Trabalho: Fundamentos para a Gestão de Pessoas* (pp. 159–176). Porto Alegre: Artmed.
- Pilati, R. & Borges-Andrade, J. E. (2006). Construção de medidas e delineamentos em avaliação de TD&E. In J. E. Borges-Andrade, G. da S. Abbad, & L. Mourão (Eds.), *Treinamento, Desenvolvimento e Educação em Organizações e Trabalho: Fundamentos para a Gestão de Pessoas* (pp. 359–384). Porto Alegre: Artmed.
- Pires, A. K., Prata, C. F., Santos, D. M. dos, Brandão, H. P., Moraes, H., Carvalho, I. M. de, ... Dantas, V. C. (2005). *Gestão por competências em organizações do governo: mesa redonda de pesquisa-ação*. Brasília: ENAP.
- Renner, D., Laumer, S., & Weitzel, T. (2014). Effectiveness and efficiency of blended learning—A literature review.
- Ribeiro, L. C. (2010). *Impactos da anualidade orçamentária na alocação de recursos públicos*. Secretaria de Orçamento Federal – SOF.
- Rosano-Peña, C. (2008). Um Modelo de Avaliação da Eficiência da Administração Pública através do Método Análise Envoltória de Dados (DEA). *Revista de Administração Contemporânea*, 12(1), 83–106.
- Rousseau, D., Wilby, J., Billingham, J., & Blachfellner, S. (2018). *Scientific Principles for General Systemology* (Vol. 13). Singapore: Springer Singapore
- Sander, J. R. (2003). Introduction. In: Kellaghan, T., & Stufflebeam, D. L. (Eds.). *International Handbook of Educational Evaluation: Part One: Perspectives/Part Two: Practice* (Vol. 9). Springer Science & Business Media.
- Shaw, I., Shaw, I. G. R., Greene, J. C., & Mark, M. M. (Eds.). (2006). *The Sage handbook of evaluation*. Sage.
- Stewart, J. et al. (2010). Strategic HRD and the learning and development function. In: Gold, J. at al.; Holden, R.; Iles, P.; Stewart, J.; Beardwell, J. *Human Resource Development: Theory and Practice*. 1ª Ed. Reino Unido: Palgrave & Mcmillan.
- Stufflebeam, D. L., & Coryn, C. L. (2014). *Evaluation theory, models, and applications* (Vol. 50). John Wiley & Sons.
- Salas, E., & Cannon-Bowers, J. A. (2001). The Science of Training: A Decade of Progress. *Annual Review of Psychology*, 52(1), 471–499. <https://doi.org/10.1146/annurev.psych.52.1.471>

- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The Science of Training and Development in Organizations. *Psychological Science in the Public Interest*, 13(2), 74–101. <https://doi.org/10.1177/1529100612436661>
- Salcedo, R. E. (2020). Performance efficiency of the Teacher Education Programs of a state university in the Philippines: A data envelopment analysis study. *Journal of Critical Reviews*, 7(7), 96–103. <https://doi.org/10.31838/jcr.07.07.16>
- Sant'Anna, T. D., Lopes, A. V., Miranda, R. F. D. A., Bermejo, P. H. D. S., & Demo, G. (2020). Scientific Research on the Efficiency of Public Expenditures: How and Where is it Going? *International Journal of Public Administration*, 43(11), 926–936. <https://doi.org/10.1080/01900692.2019.1660995>
- Scott, W. R. (1998). *Organizations: rational, natural, and open systems*. (4th ed.). Prentice Hall.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage Publications, Inc.
- Senra, L. F. A. D. C., Nanci, L. C., Mello, J. C. C. B. S. de, & Meza, L. A. (2007). Estudo sobre métodos de seleção de variáveis em DEA. *Pesquisa Operacional*, 27(2), 191–207. <https://doi.org/10.1590/S0101-74382007000200001>
- Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2012). *Metodologia de pesquisa em psicologia*. AMGH Editora.
- Siqueira, L., Farias, R., Bermejo, P. H., & Souza, W. (2020). Análise de variáveis para mensuração da eficiência do gasto público por função de governo. *Administração Pública e Gestão Social*, 12(1), 18. <https://doi.org/10.21118/apgs.v12i1.5632>
- Silva, A. L. (2011). *Avaliação de Treinamentos nos Níveis de Impacto no Trabalho e Resultados Organizacionais*. (Tese de doutorado). Universidade de Brasília, Brasil.
- Silva, A. L., & Abbad, G. da S. (2011). Benefícios e limitações do balanced scorecard para avaliação de resultados organizacionais em treinamento, desenvolvimento e educação. *Revista ibero-americana de estratégia*, 10(1), 4–28. <https://doi.org/10.5585/riae.v10i1.1695>
- Silva, F. dos S. J. da, Diniz, J. A., & Roratto, R. (2016). Qualidade do gasto e mensuração da eficiência dos cursos de capacitação para servidores públicos. *Cadernos de Finanças Públicas*, 16, 7–29.
- Smith, P. C., & Street, A. (2005). Measuring the Efficiency of Public Services: The Limits of Analysis. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 168(2), 401–417. Retrieved from <http://www.jstor.org/stable/3559969>
- Soheilrad, S., Govindan, K., Mardani, A., Zavadskas, E. K., Nilashi, M., & Zakuan, N. (2018). Application of data envelopment analysis models in supply chain management: A systematic review and meta-analysis. *Annals of Operations Research*, 271(2), 915–969.

- Song, M., An, Q., Zhang, W., Wang, Z., & Wu, J. (2012). Environmental efficiency evaluation based on data envelopment analysis: A review. *Renewable and Sustainable Energy Reviews*, 16(7), 4465-4469.
- Sowlati, T. (2005). Efficiency studies in forestry using data envelopment analysis. *Forest products journal*, 55(1), 49.
- Stegmaier, J. (2012). Effects of Works Councils on Firm-Provided Further Training in Germany. *BRITISH JOURNAL OF INDUSTRIAL RELATIONS*, 50(4), 667–689.
- Stufflebeam, D. L. (2003). The CIPP Model for Evaluation. In: Kellaghan, T., & Stufflebeam, D. L. (Eds.). *International Handbook of Educational Evaluation: Part One: Perspectives/Part Two: Practice* (Vol. 9). Springer Science & Business Media.
- Suchman, E.A. (1967). *Evaluative research: Principles and practice in public service and social action programs*. New York: Russell Sage Foundation.
- Suescún, O. Y., Cubillos, A. A., & García, L. (2017). Análisis envolvente de datos para la medición de la eficiencia en instituciones de educación superior: una revisión del estado del arte TT - Analyse entourent de données pour la mesure de l'efficacité dans les institutions d'éducation supérieure: une rev. *Revista Científica General José María Córdova*, 15(19), 147–173. <https://doi.org/10.21830/19006586.84>
- Talukder, D. (2011). Are Private Providers more Productive and Efficient than Public Providers of International Education? Evidence from New Zealand. In Annual conference of the New Zealand Association of Economists, Wellington.
- Tannenbaum, S. I., & Yukl, G. (1992). Training and Development in Work Organizations. *Annual Review of Psychology*, 43, 399–441. <https://doi.org/10.1177/103841118902700106>
- Tasca, J. E., Ensslin, L., Ensslin, S. R., & Alves, M. B. M. (2010). An approach for selecting a theoretical framework for the evaluation of training programs. *Journal of European Industrial Training*, 34(7), 631–655. <https://doi.org/10.1108/03090591011070761>
- TCU. (2013). *Levantamento de governança e gestão de pessoas: Tribunal de Contas da União. Relator Ministro-substituto Marcos Bemquerer Costa*. Brasília: TCU, Secretaria de Fiscalização de Pessoal,.
- TCU. (2016). *Levantamento de governança e gestão de pessoas: Tribunal de Contas da União. Relator Ministro Augusto Nardes* (1st ed.). Brasília: TCU, Secretaria de Fiscalização de Pessoal.
- Tharenou, P. (2010). Training and Development in Organizations. In A. Wilkinson, N. Bacon, T. Redman, & S. A. Snell (Eds.), *The Sage Handbook of Human Resource Management* (pp. 155–172). London: Sage Publications.
- Tsaples, G., & Papathanasiou, J. (2021). Data envelopment analysis and the concept of sustainability: A review and analysis of the literature. *Renewable and Sustainable Energy Reviews*, 138, 110664.

- Vargas, M.R.M., & Abbad, G.S. (2006). Bases conceituais em treinamento, desenvolvimento e educação – TD&E. In J.E. Borges-Andrade, G. Abbad, & L. Mourão (Orgs.), *Treinamento, desenvolvimento e educação em organizações e trabalho: fundamentos para a gestão de pessoas* (pp. 137–158). Porto Alegre: Artmed.
- Warr, P., Birdi, M., & Rackham, N. (1970). Evaluation of management training. 6^a.
- Wexley, K. N. (1984). Personnel training. *Annual Review of Psychology*, 35, 519–551. <https://doi.org/10.1146/annurev.ps.35.020184.002511>
- Yang, G. L., Fukuyama, H., & Song, Y. Y. (2018). Measuring the inefficiency of Chinese research universities based on a two-stage network DEA model. *Journal of Informetrics*, 12(1), 10-30.
- Zhou, H., Yang, Y., Chen, Y., & Zhu, J. (2018). Data envelopment analysis application in sustainability: The origins, development and future directions. *European Journal of Operational Research*, 264(1), 1-16.

APÊNDICE A - Lista de artigos da Revisão de Literatura

N	AUTORIA E ANO	TÍTULO	PERIÓDICO	CIT.	SCORE InOrd
1	Drost et al., (2002)	Benchmarking training and development practices: a multi-country comparative analysis	Human Resource Management	110	102,0016
2	Kong & Fu (2012)	Assessing the performance of business colleges in Taiwan using data envelopment analysis and student based value-added performance indicators	Omega: International Journal of Management Science	84	86,00258
3	Freitas & Borges-Andrade (2004)	Efeitos de treinamento nos desempenhos individual e organizacional	Revista de Administração de Empresas – RAE	79	73,00022
4	Mourão & Borges-Andrade (2005)	Avaliação de programas públicos de treinamento: um estudo sobre o impacto no trabalho e na geração de emprego	Organizações e Sociedade	70	65
5	Stegmaier (2012)	Effects of Works Councils on Firm-Provided Further Training in Germany	British Journal of Industrial Relations	47	49,00116
6	Mehra et al. (2014)	Estimating Returns to Training in the Knowledge Economy	MIS Quarterly	41	45,00453
7	Haksever & Muragishi (1998)	Measuring Value in MBA Programmes	Education Economics	52	40,00054
8	Campos et al. (2004)	Avaliação do sistema de treinamento e desenvolvimento em empresas paulistas de médio e grande porte	Psicologia: Reflexão e Crítica	43	37,00033
9	Aviles-Sacoto et al. (2015)	Two-stage network DEA: when intermediate measures can be treated as outputs from the second stage	The Journal of the Operational Research Society	22	27,00095
10	Parshakov & Shakina (2018)	With or without CU	Journal of Intellectual Capital	14	22,00118
11	Debnath & Shankar (2009)	Assessing performance of management institutions: An application of data envelopment analysis	TQM Journal	23	22,00066
12	Fisher et al. (2007)	A Value-Added Approach to Selecting the Best Master of Business Administration (MBA) Program	Journal of Education for Business	25	22,00044
13	Hirao (2012)	Efficiency of the top 50 business schools in the United States	APPLIED ECONOMICS LETTERS	16	18,00037
14	Hsu et al. (2009)	An Efficiency Comparison of MBA Programs: Top 10 Versus Non-Top 10	Journal of Education for Business	18	17,00044
15	Meneses & Abbad (2009)	Proposta para Desenvolvimento de Modelos de Avaliação da Efetividade de Programas de Treinamento.	Revista de Administração Contemporânea - RAC	18	17
16	Halachev (2012)	Prediction of e-learning efficiency by neural networks	Cybernetics and Information Technologies	14	16,00031

17	Ekiz (2020)	A new DEA approach to fully rank DMUs with an application to MBA programs	INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH	4	14,00102
18	Blinov et al. (2019)	MANAGEMENT OF REGIONAL TRAINING SYSTEMS	The Education and science journal	5	14
19	Kumar (2019)	Artificial intelligence divulges effective tactics of top management institutes of India	BENCHMARKING-AN INTERNATIONAL JOURNAL	4	13,00055
20	Gagnon & Smith (2013)	The Effects of a Training Levy on Training Characteristics and Outcomes: The Case of Quebec	Relations Industrielles	10	13,00038
21	Baxter (2014)	Undergraduate interpreter training in the Spanish state: An analytical comparison	Sendeban	9	13,00015
22	Boudet et al. (2017)	Evaluación del impacto de la capacitación en directivos de instalaciones turísticas	Ciencia, docencia y tecnología	6	13
23	Teixeira Filho et al. (2017)	Capacitação no Setor Público: Analisando o Processo de uma Ifes	RACE: Revista de Administração, Contabilidade e Economia	6	13
24	Barrera (2017)	Relación de Capacitación con Adopción de Internet y E-Commerce: Diferencias entre Microemprendedores de Chile	Información tecnológica	5	12,00023
25	Franco et al. (2015)	A andragogia na educação corporativa: o caso de uma empresa metalúrgica	Revista de Administração da Unimep	7	12
26	Othman & Naintin (2016)	The relationship between maritime education and employer trust: the structural equation modelling (SEM) perspective	WMU JOURNAL OF MARITIME AFFAIRS	5	11,0004
27	Sapper & Coronel (2016)	A Percepção dos Servidores Técnico-Administrativos da UFSM, Campus de Frederico Westphalen, Quanto à Efetividade dos Cursos de Capacitação da Instituição na Atividade Laboral	Revista de Administração do IMED	5	11
28	Souza et al. (2014)	Modelos organizacionais para treinamento versus estilos de aprendizagem de colaboradores: um estudo de caso em uma empresa de serviços de infraestrutura organizacional	Revista Gestão e Planejamento	7	11
29	Nojavan et al. (2020)	A fuzzy service quality based approach for performance evaluation of educational units	Socio-Economic Planning Sciences	0	10,00132
30	Osorio-Londoño et al. (2020)	Training and its influence on competitive strategy implementation	Human Resource Development Quarterly	0	10,00086
31	Meira & Bandeira (2018)	Literacy in the cooperative sector evaluated through the reserve for cooperative education and training	CIRIEC-Espana Revista de Economía Publica, Social y Cooperativa	2	10,00033
32	Fonina et al. (2019)	The conceptual foundations of educational cooperation of universities and companies of Industry 4.0	ON THE HORIZON	1	10,00017
33	Nkonki et al. (2014)	Efficiency in foundation provisioning in a selected university	Mediterranean Journal of Social Sciences	6	10,00014

34	Salcedo (2020)	Performance efficiency of the Teacher Education Programs of a state university in the Philippines: A data envelopment analysis study	Journal of Critical Reviews	0	10
35	Guzman-Anaya (2019)	Knowledge Transfer in the Automotive Industry: The Case of JICA's Project for Automotive Supply Chain Development in Mexico	México y la cuenca del Pacífico	1	10
36	Fisher et al. (2017)	Determining the Value of Undergraduate Business Programs from Market vs Academic Perspectives	International Journal of Educational Management	2	9,000567
37	Damasceno et al. (2012)	Modelos lógicos e avaliações de treinamentos organizacionais	Paideia	7	9,000195
38	Ankudinov et al. (2018)	Empirical Analysis of the Relation Between Expenditures on Employees' Advanced Training and Financial Performance of Russian Companies	UPRAVLENETS-THE MANAGER	1	9
39	Pai et al. (2019)	Data Envelopment Analysis (DEA) - Application at NMIMS - SBM, a Leading AACSB Accredited Indian Higher Education Business School	NMIMS MANAGEMENT REVIEW	0	9
40	Ensslin et al. (2018)	Gestão Pública: Construção de um Modelo Construtivista para Apoiar a Escola Virtual de Administração Pública de Santa Catarina no Combate à Evasão da Capacitação a Distância	Revista Ibero-Americana de Estratégia	1	9
41	Bogetoft & Wittrup (2017)	Apples to oranges: benchmarking vocational education and training programmes	SCHOOL EFFECTIVENESS AND SCHOOL IMPROVEMENT	1	8,001145
42	Kao, Ling-Jing; Chiu, Shu-Yu; Ko, Hsien-Tang	A Study of the Talent Training Project Management for Semiconductor Industry in Taiwan: The Application of a Hybrid Data Envelopment Analysis Approach	SCIENTIFIC WORLD JOURNAL	4	8,000384
43	Satybaldiyeva et al. (2018)	Determination of effective balanced indicators in the airline company using a modified fuzzy analytical hierarchy process approach	Journal of Advanced Research in Law and Economics	0	8,000229
44	Borja Lira et al. (2009)	Opinión de los docentes sobre el programa de capacitación y su relación con el desempeño escolar en Sonora, 2004-2009	Región y sociedad	9	8
45	Carvalho & Chiaretto (2018)	A importância da educação continuada para o serviço operacional no corpo de bombeiros de Minas Gerais	Revista Metropolitana de Governança Corporativa	0	8
46	Silva et al. (2016)	Qualidade do gasto e mensuração da eficiência dos cursos de capacitação para servidores públicos.	Cadernos de Finanças Públicas	1	7
47	Mourão & Marins (2010)	Quem está motivado para aprender nos cursos oferecidos pelas empresas?	Paideia	6	6,000195
48	Petrenko (2015)	THE EFFECT OF ADAPTATION LEVEL OF SYSTEM OF	SCIENTIFIC BULLETIN OF POLISSIA	1	6

		EMPLOYEE JOB TRAINING ON ALTERNATION OF FINANCIAL INDICATORS OF ENTERPRISE			
49	Lopes et al. (2016)	ROI em treinamento: aplicação de uma metodologia para mensuração	Revista Inovação, Projetos e Tecnologias	0	6
50	Jessop (2010)	A portfolio model for performance assessment: the Financial Times MBA ranking	The Journal of the Operational Research Society	5	5,000951
51	Ng & Li (2003)	Training and Productive Efficiency in Transition Economies	Oxford Economic Papers	12	5,000743
52	Brene et al. (2015)	ANALYSIS EFFICIENCY DETERMINANTS IN ENTERPRISE SERVICES CIARAMA MACHINES PONTA PORA/MS FOR THE YEARS 2012 AND 2013	REVISTA CIENTIFICA HERMES	0	5
53	Talukder (2011)	Are Private Providers more Productive and Efficient than Public Providers of International Education? Evidence from New Zealand.	Oeconomics of Knowledge	4	5
54	Contreras & Gonzalez (2013)	Retorno do investimento em formação da pequena empresa em Cartagena	Estudios Gerenciales	0	3
55	Odelius & Siqueira Júnior (2011)	Construção e validação de instrumento de pesquisa relativo a aspectos que influenciam a efetividade de políticas de TD&E	Contabilidade, Gestão e Governança	2	3
56	Sampaio & Tavares (2001)	Estrutura e programas de T&D: o caso das empresas públicas e sociedades de economia mista do Estado de Minas Gerais	Revista de Administração Contemporânea - RAC	12	3
57	Schröder e. al. (2005)	Sistemas de Treinamento Corporativo virtual: definindo critérios e indicadores de avaliação	Revista Eletrônica de Administração - REAd.	8	3
58	Cubbin & Zamani (1996)	A comparison of performance indicators for training and enterprise councils in the UK	Annals of Public and Cooperative Economics	16	2,00045
59	Debnath et al. (2008)	Performance Evaluation of Technical Institutions: An Application of Data Envelopment Analysis	Education, Knowledge & Economy: A Journal for Education and Social Enterprise	2	0
60	Martinez & Olarra (2010)	Capacitar para competir	Revista ADM.MADE	0	0
61	Silva & Oishi (2000)	Um processo de avaliação do efeito do treinamento para melhoria da lucratividade nas empresas: uma avaliação conjugada de produtividade física e faturamento	Revista de Negócios	0	-10

APÊNDICE B – Carta-convite

Assunto: Solicitação de Apoio à Pesquisa

De: Diogo Ribeiro da Fonseca

A/C: Gestores(as) de Pessoas e Coordenadores(as) de Desenvolvimento de Pessoas nas Organizações Públicas.

Projeto de Pesquisa: Desenvolvimento de metodologia para Avaliação da Política Nacional de Desenvolvimento de Pessoas (PNDP) no período 2015-2020.

Prezado(a) Gestor(a),

O presente documento tem como objetivo convidar a V. Sa. e sua organização para participação junto à pesquisa para a Tese de Doutorado em Administração do aluno Diogo Ribeiro da Fonseca da Universidade de Brasília, sob a orientação do Prof. Doutor Pedro Paulo Murce Meneses do Programa de Pós-Graduação em Administração.

Tendo em vista a nova regulamentação da Política Nacional de Desenvolvimento de Pessoas – PNDP por meio do Decreto n.º 9.991/2019, em substituição ao Decreto n.º 5.707/2006, a presente pesquisa se propõe a realizar uma avaliação retrospectiva de como se deram as políticas de capacitação em variadas organizações públicas federais nos últimos cinco anos, englobando o período de 2015 a 2019. Para tanto, a pesquisa visa coletar dados com o objetivo de avaliar o **alinhamento estratégico**, o **custo-eficácia** e a **abrangência** dos cursos ofertados nas organizações, tendo como referência os objetivos da nova PNDP.

Sendo assim, a pesquisa visa contribuir com o desenvolvimento de um método de avaliação que possibilite apontar estratégias de gestão e avaliação para melhoria das políticas organizacionais de capacitação no setor público. Com esse objetivo, solicitamos a colaboração de V. Sa. de prover os dados abaixo relacionados por meio do envio de documentos e preenchimento de formulários que visam levantar os fatores relevantes para a oferta de cursos pelas unidades de capacitação aos seus servidores públicos em suas organizações.

Dados solicitados

A principal informação que solicitamos para a presente pesquisa refere-se ao fornecimento das **planilhas de controle** relativas à execução dos programas e ações de capacitação e desenvolvimento efetivados em sua organização para cada um dos anos de 2015, 2016, 2017, 2018 e 2019. Especificamente, os seguintes elementos são de suma importância para o desenvolvimento desta pesquisa:

- **Nome** da ação de capacitação (curso, palestra, evento científico, workshop, etc.) *;
- **Organização fornecedora** das ações de capacitação e desenvolvimento;
- Forma de contratação ou forma de realização do curso: com o objetivo de identificar se o curso foi realizado por fornecedores **externos** ou **internamente** pelo próprio órgão;
- **Modalidade** de ensino da ação de capacitação e desenvolvimento;
- **Tipo** da ação de capacitação e desenvolvimento;

- **Área temática** da ação de capacitação e desenvolvimento;
- **Carga horária** da ação de capacitação e desenvolvimento*;
- Número de **participantes** da ação de capacitação e desenvolvimento*;
- **Custo** da ação de capacitação e desenvolvimento*;
- Data de início ou **período** da ação de capacitação e desenvolvimento*.

* *Dados fundamentais para a análise.*

Obs.: Não é necessário tratar os dados. O tratamento dos dados será realizado pelo autor da pesquisa. Solicitamos apenas o envio da planilha original utilizada para registrar essas informações.

Exemplo de planilha:

Ação	Fornecedor	Forma de contratação	Modalidade de ensino	Tipo da ação	Área temática	Carga horária	Número de participantes	Custo total da ação	Data de início
Treinamento de lideranças	Empresa S.A	Dispensa de licitação	Presencial	Curso	Gestão de Pessoas	40h	12	R\$ 5.600,00	15/01/2014
Introdução à finanças públicas	University	n/a	à distância	Evento	Finanças	20h	50	Gratuito	02/02/2015
SIAPE Gerencial	Prof. Américo Pinheiro	GECC	Presencial	Curso	ADM. Pública	10h	30	R\$ 2.000,00	08/01/2014

Além da informação sobre as ações de capacitação realizadas, solicitamos o envio de informações acerca do **contexto** em que se encontrava a organização e a unidade de capacitação em cada um dos anos, por meio dos **documentos** a seguir:

Documentos solicitados

- 1) **Planejamentos Estratégicos** da Organização vigentes nos anos de 2015, 2016, 2017, 2018 e 2019.
- 2) **Planos Anuais de Capacitação** vigentes nos anos de 2015, 2016, 2017, 2018 e 2019.
- 3) **Políticas de Capacitação** – se for o caso – nos anos de 2015, 2016, 2017, 2018 e 2019.
- 4) Preenchimento do **Formulário de Descrição de Equipe** com informações sobre integrantes da equipe responsável pela capacitação nos anos de 2015, 2016, 2017, 2018 e 2019. *
- 5) Preenchimento do **Formulário de Descrição de Processos** com informações sobre atividades de planejamento e avaliação da capacitação realizadas nos anos de 2015, 2016, 2017, 2018 e 2019. *

* Formulários enviados em anexo no e-mail de convite.

Adicionalmente, caso deseje, podem ser enviadas outras fontes de informação tais como relatórios de capacitação, relatórios de sistemas, de resultados, etc.

Resultados Esperados

O estudo realizado nesta Tese de Doutorado constitui parte da agenda de pesquisa para a melhoria da governança da capacitação no setor público federal. Para referência acerca dessa linha de pesquisa, pode-se consultar o Caderno ENAP de volume 54 “Gestão de redes de capacitação: metodologia para um modelo de governança” no endereço: <http://repositorio.enap.gov.br/handle/1/3668> . Os resultados esperados dessa linha de pesquisa são os seguintes:

- Aperfeiçoamento do monitoramento, avaliação e planejamento das atividades de capacitação nas organizações públicas federais;
- Subsídio ao desenvolvimento de instrumentos de gestão da PNDP;

- Informatização da gestão da capacitação;
- Aumento de competitividade e redução de custos das ações de capacitação;
- Mapeamento de competências organizacionais e individuais.

Os resultados da presente pesquisa estarão disponíveis a partir de **março de 2021** e serão apresentados às organizações públicas participantes. Os dados serão utilizados para fins estritamente acadêmicos. A pesquisa visará, futuramente, avaliar, também, a capacitação no ano de 2020, haja vista os impactos decorrentes da epidemia do COVID-19. Almeja-se, assim, o monitoramento sistemático da capacitação federal, visando a sua constante melhoria.

Dados para Envio e Contato

Apresentados os motivos e os dados necessários para a pesquisa, reforçamos a importância de sua cooperação e contribuição para a concretização desse projeto. Sendo assim, solicitamos, por gentileza, o envio dos dados e informações para o e-mail: **redescapitacao@gmail.com** . Solicitamos, ainda, a disponibilização de um **contato de servidor ou colaborador** responsável com o qual possamos nos comunicar com facilidade em caso de dúvidas ou ajustes acerca dos dados e da pesquisa, tendo em vista o atual contexto de **teletrabalho**.

Em caso de dúvidas ou qualquer necessidade de esclarecimento, não hesite em entrar em contato por meio do n.º +55 (61) 98539-4325 ou pelo e-mail **diogorfonseca@gmail.com** . Por fim, tendo em vista os prazos para tratamento e análise dos dados, solicitamos, por gentileza, o envio dos documentos **até o dia 15/01/2020**.

Certos de sua colaboração.

Agradecemos desde já.

Pesquisador Responsável

Diogo Ribeiro da Fonseca, Msc.

Doutorando do Programa de Pós-Graduação em Administração da Universidade de Brasília Professor da Graduação em Administração Pública do Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa – IDP

DIOGO RIBEIRO DA FONSECA
UNIVERSIDADE DE BRASÍLIA

APÊNDICE C - Formulário de Descrição de Equipe

Projeto de Pesquisa: Desenvolvimento de metodologia para Avaliação da Política Nacional de Desenvolvimento de Pessoas (PNDP) no período 2015-2020.

Pesquisador Responsável: Diogo Ribeiro da Fonseca – Tel.: +55 (61) 98539-4325. E-mail: diogorfonseca@gmail.com

Prezado(a),

O presente formulário visa coletar dados para auxílio à pesquisa da Tese de Doutorado em Administração do aluno Diogo Ribeiro da Fonseca da Universidade de Brasília, sob a orientação do Prof. Doutor Pedro Paulo Murce Meneses. As informações coletadas através deste instrumento serão utilizadas em caráter estritamente acadêmico.

O quantitativo e o perfil da equipe constitui-se como um dos principais fatores que afetam a produtividade das unidades gerenciais no setor público. Tendo em vista esse aspecto, informe por gentileza as seguintes informações a respeito da **equipe integrante da unidade responsável pelo fornecimento de ações de capacitação em sua organização** para os anos de 2015, 2016, 2017, 2018 e 2019. **Tome como referência a situação na data aproximada de 31/12 de cada ano.**

1) Quantitativo de Pessoal e Cargos					
	31/12/2015	31/12/2016	31/12/2017	31/12/2018	31/12/2019
Total de Pessoas*					
Nº de Servidores: Nível Médio					
Nº de Servidores: Nível Superior					
Nº de Terceirizados					
Nº de Estagiários					
Qtd. de Cargos / Funções comissionadas					

2) Escolaridade					
	31/12/2015	31/12/2016	31/12/2017	31/12/2018	31/12/2019
Nº de pessoas até o Nível Médio					
Nº de pessoas com Nível Superior					
Nº de pessoas com Especialização					
Nº de pessoas com Mestrado					
Nº de pessoas com Doutorado					

* O único dado obrigatório para a análise é o “Total de Pessoas” em cada ano. Caso não possua os demais dados, não hesite em enviar o instrumento incompleto ou com dados parciais.

A sua contribuição é de grande importância para este trabalho e recebida com muita gratidão.

APÊNDICE D – Formulário de Descrição de Processos de Capacitação

Projeto de Pesquisa: Desenvolvimento de metodologia para Avaliação da Política Nacional de Desenvolvimento de Pessoas (PNDP) no período 2015-2020.

Pesquisador Responsável: Diogo Ribeiro da Fonseca – Tel.: +55 (61) 98539-4325. E-mail: diogorfonseca@gmail.com

Prezado(a),

O presente formulário visa coletar dados para auxílio à pesquisa da Tese de Doutorado em Administração do aluno Diogo Ribeiro da Fonseca da Universidade de Brasília, sob a orientação do Prof. Doutor Pedro Paulo Murce Meneses. As informações coletadas através deste instrumento serão utilizadas em caráter estritamente acadêmico.

O processo de Treinamento é constituído de elementos e etapas técnicas que podem fornecer suporte para a adequada gestão da capacitação. Tendo em vista esse aspecto, informe por gentileza se foram **executadas as seguintes atividades listadas abaixo pela unidade de capacitação em sua organização** nos anos de 2015, 2016, 2017, 2018 e 2019. Posteriormente informe a respeito da vinculação estratégica e existência de sistema informatizado na capacitação. **Indique com um X** nos anos correspondentes em que a atividade foi realizada. Deixe em branco caso a atividade não tenha sido realizada. **Utilize o campo “Observação” (Obs.) caso deseje.**

3) Avaliação de Necessidades de Treinamento

Descrição: A avaliação - análise, ou levantamento - de necessidades de treinamento consiste em uma pesquisa com as unidades organizacionais para identificar as suas principais necessidades de capacitação. A pesquisa pode ser feita por meio de formulários abertos, questionários, sistemas ou até mesmo entrevistas de consulta às áreas.

	2015	2016	2017	2018	2019
Essa atividade foi realizada nos anos:					

Obs:

4) Mapeamento de Competências

Descrição: O mapeamento de competências consiste em um modelo de descrição dos principais conhecimentos, habilidades e atitudes a serem desenvolvidos por servidores públicos para o desempenho de suas atividades, o qual serve de referência para o planejamento dos treinamentos realizados pela organização.

	2015	2016	2017	2018	2019
Essa atividade foi realizada (ou estava vigente) nos anos:					

Obs:

5) Plano de Capacitação

Descrição: Os planos de capacitação descrevem os principais cursos a serem realizados pela organização no(s) próximo(s) exercício(s), podendo ser detalhados por tipos de cursos, áreas de conhecimento, unidades organizacionais, competências e distribuição dos recursos de capacitação.

	2015	2016	2017	2018	2019
Essa atividade foi realizada nos anos:					

Obs:

6) Avaliação de Reação

Descrição: A unidade de capacitação pode realizar diferentes tipos de avaliação de seus treinamentos, conforme seus objetivos. A Avaliação de Reação ao Treinamento verifica a qualidade da programação, instrutores, materiais fornecidos, entre outros aspectos do treinamento realizado.

	2015	2016	2017	2018	2019
Essa atividade foi realizada nos anos:					

Obs:

7) Avaliação de Aprendizagem

A Avaliação de Aprendizagem verifica, formalmente, se os alunos apreenderam o conteúdo lecionado.

	2015	2016	2017	2018	2019
Essa atividade foi realizada nos anos:					

Obs:

8) Avaliação de Impacto

A Avaliação de Impacto de Treinamento verifica, após um determinado período de tempo, quais resultados o treinamento gerou para o desempenho do aluno em seu contexto de trabalho.

	2015	2016	2017	2018	2019
Essa atividade foi realizada nos anos:					

Obs:

Por favor, informe se a sua organização possuía um Plano Estratégico nos anos listados e se este incluiu metas e objetivos para a capacitação na organização:

9) Plano Estratégico Organizacional					
	2015	2016	2017	2018	2019
A Organização possuía um plano estratégico vigente nos anos:					
A Organização possuía um plano estratégico com objetivos e metas de capacitação nos anos:					
Obs:					

Por favor, informe se a sua organização possuía um Sistema Informatizado para a gestão da capacitação

10) Sistema de Capacitação					
	2015	2016	2017	2018	2019
A Organização possuía um sistema informatizado de capacitação nos anos:					
Obs:					

A sua contribuição é de grande importância para este trabalho e recebida com muita gratidão.

APÊNDICE E – Base de dados final

DMU	Insumo e Produtos				Carga Horária Absoluta e Proporcional por Tipo e Modalidade de Ensino						Servidores		
	(X1) Equipe	(X2) Despesas	(Z) Carga Horária	(Y) Participações	Presencial Externa		Presencial Interna		EAD		Servidores Capacitados	Força de Trabalho	% Capacitado
A5	6	R\$ 468.519,34	6.143	2643	3.957	64,4%	264	4,3%	1.922	31,3%	641	1310	48,9%
A6	7	R\$ 420.189,08	5.852	2294	2.930	50,1%	420	7,2%	2.502	42,8%	472	1310	36,0%
A7	7	R\$ 560.260,28	6.642	1988	3.969	59,8%	266	4,0%	2.407	36,2%	404	1263	32,0%
A8	9	R\$ 1.151.019,56	3.990	2214	2.456	61,6%	371	9,3%	1.163	29,1%	690	1237	55,8%
A9	10	R\$ 608.249,91	5.285	2504	2.135	40,4%	334	6,3%	2.816	53,3%	796	1189	67,0%
B5	4	R\$ 363.309,08	3.780	732	3.077	81,4%	483	12,8%	220	5,8%	485	2056	23,6%
B6	4	R\$ 543.155,74	4.065	360	3.672	90,3%	210	5,2%	183	4,5%	277	2071	13,4%
B7	4	R\$ 463.441,97	5.214	361	4.730	90,7%	224	4,3%	260	5,0%	297	1895	15,7%
B8	5	R\$ 555.547,87	4.090	247	3.569	87,3%	165	4,0%	356	8,7%	221	1862	11,9%
B9	5	R\$ 173.533,45	1.273	141	1.163	91,4%	80	6,3%	30	2,4%	134	1726	7,8%
C5	13	R\$ 1.163.696,07	9.540	3036	7.561	79,3%	427	4,5%	1.552	16,3%	634	730	86,8%
C6	13	R\$ 577.441,01	12.305	2707	8.328	67,7%	810	6,6%	3.167	25,7%	600	715	83,9%
C7	12	R\$ 479.808,00	7.324	4510	4.712	64,3%	702	9,6%	1.910	26,1%	663	706	93,9%
C8	10	R\$ 1.081.035,36	10.177	2872	6.310	62,0%	257	2,5%	3.610	35,5%	648	695	93,2%
C9	9	R\$ 741.841,73	10.546	2875	5.704	54,1%	441	4,2%	4.401	41,7%	662	679	97,5%
D5	8	R\$ 1.087.947,59	7.153	2379	6.126	85,6%	443	6,2%	584	8,2%	714	854	83,6%
D6	7	R\$ 639.785,16	7.121	2239	5.499	77,2%	271	3,8%	1.351	19,0%	683	838	81,5%
D7	7	R\$ 658.862,97	7.018	1673	5.116	72,9%	391	5,6%	1.511	21,5%	619	838	73,9%
D8	6	R\$ 2.622.944,43	7.843	2585	4.607	58,7%	1.230	15,7%	2.006	25,6%	685	824	83,1%
D9	6	R\$ 1.028.733,82	7.169	2160	5.954	83,1%	377	5,3%	838	11,7%	622	855	72,7%

DMU	Insumo e Produtos				Carga Horária Absoluta e Proporcional por Tipo e Modalidade de Ensino						Servidores		
	(X1) Equipe	(X2) Despesas	(Z) Carga Horária	(Y) Participações	Presencial Externa		Presencial Interna		EAD		Servidores Capacitados	Força de Trabalho	% Capacitado
E5	3	R\$ 47.820,00	873	58	873	100,0%	-	0,0%	1	0,1%	44	1006	4,4%
E6	3	R\$ 24.190,00	562	37	410	73,0%	-	0,0%	152	27,0%	26	997	2,6%
E7	3	R\$ 54.453,60	958	161	557	58,1%	111	11,6%	290	30,3%	135	938	14,4%
E8	3	R\$ 168.427,00	286	186	226	79,0%	60	21,0%	1	0,3%	117	886	13,2%
E9	3	R\$ 568.303,10	826	326	798	96,6%	28	3,4%	1	0,1%	221	780	28,3%
F5	6	R\$ 1.083.284,15	3.036	2808	1.280	42,2%	256	8,4%	1.500	49,4%	1997	3297	60,6%
F6	6	R\$ 787.210,30	1.657	2690	979	59,1%	153	9,2%	525	31,7%	2011	3077	65,4%
F7	6	R\$ 758.105,98	3.224	1807	1.384	42,9%	123	3,8%	1.717	53,3%	1364	3038	44,9%
F8	7	R\$ 312.577,69	1.533	958	1.261	82,3%	256	16,7%	16	1,0%	595	2871	20,7%
F9	7	R\$ 433.888,83	1.192	768	1.118	93,8%	54	4,5%	20	1,7%	571	2726	20,9%

APÊNDICE F – Base de dados: Fatores não discricionários dicotomizados

DMU	Variáveis Dicotomizadas do Processo de Treinamento								Variáveis Dicotomizadas da Estrutura Organizacional		
	ANT	Map. De Compet	Plano de Capacit.	Av. de Reação	Av. de Aprend.	Av. de Impacto	Estratégia e Objetivos de Capacit.	Sistema de Capacit.	Possui Regionais Codificação	Nível Hierárquico Codificação	Porte (+1.100)
A5	1	-	1	1	-	-	1	-	1	-	1
A6	1	-	1	1	-	-	1	-	1	-	1
A7	1	-	1	1	-	-	1	-	1	-	1
A8	1	-	1	1	-	-	1	-	1	-	1
A9	1	-	1	1	-	-	1	-	1	-	1
B5	-	-	-	-	-	-	-	-	1	1	1
B6	1	1	1	-	-	-	1	-	1	1	1
B7	1	1	-	-	-	-	1	-	1	1	1
B8	1	1	1	-	-	-	1	-	1	1	1
B9	1	-	1	-	-	-	1	-	1	1	1
C5	1	1	1	1	-	-	1	-	-	1	-
C6	1	1	1	1	-	-	1	-	-	1	-
C7	1	1	1	1	-	-	1	-	-	1	-
C8	1	1	1	1	-	-	1	-	-	1	-
C9	1	1	1	1	-	-	1	-	-	1	-
D5	1	-	1	1	1	-	1	1	-	1	-
D6	1	-	1	1	1	1	1	1	-	1	-
D7	1	1	1	1	1	-	1	1	-	1	-
D8	1	1	1	1	1	1	1	1	-	1	-
D9	1	1	1	1	1	-	1	1	-	1	-

DMU	Variáveis Dicotomizadas do Processo de Treinamento								Variáveis Dicotomizadas da Estrutura Organizacional		
	ANT	Map. De Compet	Plano de Capacit.	Av. de Reação	Av. de Aprend.	Av. de Impacto	Estratégia e Objetivos de Capacit.	Sistema de Capacit.	Possui Regionais Codificação	Nível Hierárquico Codificação	Porte (+1.100)
E5	1	-	1	1	-	-	-	-	1	-	-
E6	-	-	1	1	-	-	-	-	1	-	-
E7	1	-	-	1	-	-	-	-	1	-	-
E8	-	-	1	1	-	-	-	-	1	1	-
E9	-	-	1	1	-	-	1	-	1	1	-
F5	1	-	1	1	-	-	-	-	-	-	1
F6	1	-	1	1	-	-	-	-	-	-	1
F7	1	-	1	1	-	-	-	-	-	-	1
F8	1	-	1	1	-	-	-	-	-	-	1
F9	1	-	1	1	1	1	-	-	-	-	1

APÊNDICE G - *Script* final no software R

ANÁLISE DE EFICIÊNCIA DE UNIDADES DE TREINAMENTO

INTRODUÇÃO: ESCOPO E ROTEIRO DE ANÁLISE

Essa análise avalia a eficiência dos processos de provisão de treinamentos a servidores em unidades de capacitação do serviço público. A amostra se constitui de 6 unidades de capacitação em órgãos públicos ao longo de um período de 05 anos (06x05). Cada organização em cada ano é considerado um caso (Decision Making Unit - DMU) independente para constituição de uma amostra relevante que permita distinguir ao longo dos cinco anos, quais as ofertas mais eficientes de treinamento, constituindo-se, portanto, em uma análise de 30 casos em análise intertemporal (Ferreira & Gomes, 2020).

#O mencionado processo será analisado considerando-se duas etapas da provisão de treinamento: 1) Produção e 2) Entrega. A primeira etapa analisa a eficiência da conversão de insumos [Equipe (X1)] e [Recursos Financeiros (X2)] em horas de capacitação [Carga Horária (Z)].

#São considerados treinamentos produzidos internamente, ou contratados externamente, tanto presenciais quanto EAD. Estão excluídas ações de educação (graduação, pós, etc.) assim como cursos de idiomas e incentivos educacionais por ressarcimento por serem ações distintas de Desenvolvimento e Educação. A razão para exclusão é propiciar a homogeneidade dos produtos analisados (cursos) e similaridade dos processos produtivos entre as organizações, um pressuposto da análise DEA.

#A segunda etapa analisa a eficiência do conversão e distribuição da Carga Horária contratada (Z) no máximo de oportunidades de treinamento [Participações (Y1)]. Subsidiariamente, se avaliará a proporção da força de trabalho treinada [Indivíduos Capacitados/Força de Trabalho (Y2)]. Os dois produtos Y são correlacionados, sendo que apenas o primeiro por ser um valor absoluto será considerado na definição do índice de eficiência técnica. Valores absolutos podem ser preferíveis a fatores relativos percentuais por não terem um limite superior (100%) se ajustando melhor à análise de produtividade, além de representarem diretamente o que foi produzido, ou o desempenho factual da unidade.

RESUMO DO MODELO DE ANÁLISE APLICADO

Sumariamente, esse roteiro operacionaliza a seguinte variante de análise DEA: DEA em rede de duas etapas com Análise Intertemporal e de Variáveis não-discricionárias.

Constituindo-se de três linhas metodológicas de análise combinadas:

- # 1) DEA em rede de duas etapas (Two-Stage Network DEA)
- # 2) Análise intertemporal de Dados em Painel (Panel Data)
- # 3) Análise de Variáveis não-discricionárias (Non-discretionary variables)

As "etapas" representam a divisão do processo analisado sendo aplicado o NetWork-DEA (N-DEA) em que se avalia cada etapa separadamente. Nesse modelo a primeira etapa consiste em um grupo de insumos (X) gerando um grupo de produtos (Z). Os produtos (Z) da primeira etapa servem de insumo para segunda etapa, gerando o grupo de produtos finais (Y). Os produtos intermediários (aqui chamados de "Z") interligam as duas etapas, formando a Rede (Network) de processos.

#A eficiência de cada etapa é analisada separadamente e posteriormente cria-se um índice agregado do processo integral por meio da multiplicação dos índices de eficiência das duas etapas. Uma organização deve ser eficiente nas duas etapas para ser eficiente no processo. Essa é uma das técnicas básicas para criação de um índice agregado em um N-DEA, sendo que existem outras possíveis técnicas visando considerar outros pressupostos e condicionantes para otimização da produtividade. O presente caso utilizará o modelo multiplicativo por ser o mais amplamente adotado e com pressupostos básicos para processos produtivos simples e com poucas variáveis.

RESUMO DO ROTEIRO DE ANÁLISE

O Primeiro estágio de análise da eficiência construirá uma curva (fronteira) de eficiência (InsumoxProduto) única para os 30 casos. Os casos são comparados com relação a essa fronteira criando-se um ranking de eficiência. As análises realizadas são as seguintes:

1) Índices de eficiência para todos os pressupostos:

a) Retornos Fixos e Orientação para Insumo (input): Os investimentos em cursos geram retornos constantes e a eficiência é dada pela economia de recursos [Etapa 1] e horas/aula [Etapa 2].

b) Retornos Variáveis e Orientação para Insumo (input): Os investimento em cursos geram retornos que variam (e.g. ganhos/perdas de escala) e a eficiência é dada pela economia de insumos.

c) Retornos Fixos e Orientação para Produto (Output): Os investimento em cursos geram retornos constantes e a eficiência é dada pela maior quantidade de horas/aula [Etapa 1] e oportunidades de capacitação ofertadas [Etapa 2].

d) Retornos variáveis e orientação para Output: Os investimento em cursos geram retornos variáveis e a eficiência é dada pela maior quantidade de produtos de treinamento ofertados (Intensidade do Treinamento).

2) Análise Bootstrap para definição de intervalos de confiança para os índices de eficiência em todos os casos

3) Comparação via Kolmogorov-Smirnoff (KS) dos índices para avaliação de qual dos pressupostos melhor se adequa à amostra.

4) Escolha dos pressupostos para a análise com base no teste KS e objetivos da pesquisa - provavelmente retornos variáveis com orientação para output.

4) Análise de outliers e de supereficiência para averiguar possíveis distorções oriundas de casos extremos e distinguir empresas eficientes.

4.1) Se for o caso, refazer a análise com exclusão de outliers que distorçam a análise.

5) Análise da eficiências de escala para determinação de origens e razões da ineficiência (retornos variáveis).

6) Análise de benchmarking (comparação entre pares) e de slacks (folgas) para discussão das razões da ineficiência (e.g. estratégias de ensino)

#As variáveis de Input são: Equipe (X1) e Custo (X2). As variáveis intermediárias são a carga horária de cursos. A variável critério de eficiência é Qtd. de participações. Todas as variáveis são mensuradas em número absolutos. O objetivo da análise é avaliar qual estratégia (mix de fatores) alcança a maior quantidade de oportunidades de capacitação, a qual propicia um maior % da força de trabalho capacitada. A Análise das Variáveis ambientais é realizada posteriormente, como um estudo à parte focado descrito em outro Script.

#####PREPARAÇÃO#####

#Iremos utilizar o pacote Benchmarking library(Benchmarking)

#Carregando a Planilha - O modelo selecionado é o QUARTO modelo testado e melhor ajustado à amostra, pressupostos metodológicos e objetivos de pesquisa.

library(readxl)

M4 <- read_excel("M4.xlsx")

#####ANÁLISE DA ETAPA 01#####

#Criando a matriz de insumos com a funcao cbind. "x1" e "x2".

x <- as.matrix(with(M4, cbind(X1,X2)))

separando o produto intermediário Z

z <- as.matrix(with(M4, Z))

#Verificando Outliers

```

outliers1 <- outlier.ap (x, z, NDEL = 5, NLEN = 25, TRANSPOSE = FALSE)

outlier.ap.plot(outliers1$ratio, txt=T)

outliers1$imat

#O principal outlier é o caso 19, mas vamos mantê-lo nessa etapa, avaliando se ele exibe outras características que
justifiquem sua retirada, por exemplo, superficiência ou extensão excessiva da fronteira de produção.

#Primeiramente utilizaremos o DEA tradicional - com benchmarks e folgas (slacks) - para calcular todos os índices
em todos os pressupostos:

## Retornos constantes à escala e orientação insumo (Eficiência para Retorno Constante com orientação para Input
- ECI)

eci_est01 <- dea(x,z, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)

## Retornos constantes à escala e orientação produto (Eficiência para Retorno Constante com orientação para
Output - ECO)

eco_est01 <- dea(x,z, RTS="crs", ORIENTATION = "out", SLACK=TRUE)

## Retornos variáveis à escala e orientação insumo (Eficiência para Retorno Variável com orientação para Input -
EVI)

evi_est01 <- dea(x,z, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)

## Retornos variáveis à escala e orientação produto (Eficiência para Retorno Variável com orientação para Output
- EVO)

evo_est01 <- dea(x,z, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

#Em seguida vamos rodar o DEA bootstrap para cada caso. O número recomendado de repetições é 2000 (STN,
2015)

eci_est01b <- dea.boot(x,z, NREP = 10000, RTS="crs", ORIENTATION = "in", alpha = 0.05)
eco_est01b <- dea.boot(x,z, NREP = 10000, RTS="crs", ORIENTATION = "out", alpha = 0.05)
evi_est01b <- dea.boot(x,z, NREP = 10000, RTS="vrs", ORIENTATION = "in", alpha = 0.05)
evo_est01b <- dea.boot(x,z, NREP = 10000, RTS="vrs", ORIENTATION = "out", alpha = 0.05)

#Por fim, avaliar as supereficiências para distinguir níveis de desempenho entre os eficientes.

eci_est01s <- sdea(x,z, RTS = "crs", ORIENTATION = "in")
eco_est01s <- sdea(x,z, RTS="crs", ORIENTATION = "out")
evi_est01s <- sdea(x,z, RTS="vrs", ORIENTATION = "in")
evo_est01s <- sdea(x,z, RTS="vrs", ORIENTATION = "out")

# Combinando os resultados em um banco de dados.

# CRS = Constant Returns to Scale (retornos constantes) e VRS = Variable returns to Scale (Retornos variáveis),

# in= orientação input, ou= orientação output

# Cálculo das inversas (para melhor interpretação) e cálculo das ineficiências causadas pela escala de produção
(crs/vrs).

# Do bootstrap e do sdea traremos apenas as eficiências ajustadas por essas duas técnicas.

res1 <- data.frame(crs01_in = eci_est01$eff, crs01_ou = eco_est01$eff, vrs01_in = evi_est01$eff, vrs01_ou =
evo_est01$eff, #eficiências

```

```

crs01_1ou = 1/eco_est01$eff, vrs01_1ou = 1/evo_est01$eff, #inversa

esc01_in = eci_est01$eff/evi_est01$eff, esc01_ou = eco_est01$eff/evo_est01$eff, #eficiência de
escala de in e out

crs01_inb = eci_est01b$eff.bc, crs01_oub = eco_est01b$eff.bc, #bootstrap
vrs01_inb = evi_est01b$eff.bc, vrs01_oub = evo_est01b$eff.bc, #bootstrap

scrs01_in = eci_est01s$eff, scrs01_ou = eco_est01s$eff, svrs01_in = evi_est01s$eff, svrs01_ou =
evo_est01s$eff) #Superficiências

#Exportando o Banco de Dados para Excel (Resultado 01)

library("writexl")

write_xlsx(res1,"Res1.xlsx")

#####Verificando suposição de retornos à escala com teste estatístico

#Para prosseguir com a análise e definição da Fronteira de eficiência é preciso definir quais pressupostos serão
utilizados. Com relação à orientação se utilizará, em razão do problema de pesquisa, a orientação output. As
Unidades de capacitação no setor público federal recebem um orçamento fixo anual para ser dispendido da melhor
forma com treinamento. Assim, o que se busca é a otimização dos outputs, sendo essa a orientação adotada.

#Com relação aos retornos, é preciso verificar se há diferenças nas distribuições probabilísticas dos scores de
eficiências entre retornos constantes e os índices de eficiência de retornos variáveis. Se houver diferença, isso
indica que os retornos são variáveis, a curva da fronteira de eficiência de produção é côncava e há ineficiência de
escala na amostra, devendo adotar-se como pressuposto a variabilidade dos retornos (VRS) - o retorno tende a cair
com o aumento da escala de produção.

#Teste de Kolmogorov-Smirnov - Não paramétrica para análise de igualdade entre distribuições de probabilidade
contínuas de amostras. O Teste KS normalmente é utilizado para comparação de uma distribuição com uma
distribuição Normal para se saber se a primeira atende a esse pressuposto. No presente caso, a comparação será
feita entre distribuições de resultados de índices de eficiência, visando atestar se elas são distintas.

#Ho: Não há diferença = Não há ausência de ineficiência de escala.

## (o modelo com a pressuposição de retornos constantes pode ser adotado);

#Hi: Há diferença = Há ineficiência de escala

## (o modelo com a pressuposição de retornos variáveis é o mais adequado para a análise);

#O valor crítico da estatística "D" do teste KS para uma amostra de 30 casos é de 0,24 ao p = 0,05. Se D > 0,24
rejeita-se "Ho".

#Teste dos Scores Normais para input e para output:

ks.test(eci_est01$eff, evi_est01$eff, alternative = "two.sided", exact = NULL)

ks.test(eco_est01$eff, evo_est01$eff, alternative = "two.sided", exact = NULL)

#Uma vez definida a Escala mais adequada, plotar os dados. (No caso "VRS-Output")

#####Plotando Resultados

#Isoquanta para os dois insumos

dea.plot.isoquant(M4$X1, M4$X2, RTS="vrs", txt=M4$CASO)

#Quando houver dois ou mais produtos - Função Transformação. Não é o caso.

#Função para modalidade

```

```

## Fronteira de Retornos Constantes

dea.plot.frontier(x, z, RTS="crs", txt=T)

dea.plot(x,z, RTS="crs", txt=M4$CASO, fex=0.9, pch=17)

## Fronteira de Retornos Variáveis

dea.plot(x,z, RTS="vrs", ORIENTATION = "out", txt=M4$CASO, fex=0.9, pch=17, cex.axis=0.9,
xlab="Despesas", ylab = "Horas")

dea.plot(x,z, RTS="vrs", txt=M4$CASO, fex=1, pch=17, cex.axis=0.9, xlab="Despesas", ylab = "Horas")

## Misturando as Fronteiras

dea.plot.frontier(x, z, RTS="crs", add=T)

#####Plotando intervalos de confiança oriundos do Bootstrap

#Criando um gráfico com intervalores de confiança:

# 1º Gráfico Original

dea.plot(x,z, RTS="vrs", ORIENTATION = "out", txt=T)

#Fronteira VRS

# Fronteira Corrigida

dea.plot((1/evo_est01b$eff.bc)*x,z, add = T, RTS="vrs", ORIENTATION = "out", lty = "dashed", col= "blue")

# Variação da Fronteira

dea.plot((1/evo_est01b$conf.int[,1]-evo_est01b$bias)*x,z, add = T, RTS="vrs", ORIENTATION = "out", lty =
"dotted", col = "red")

#Fronteira CRS

# Fronteira Corrigida

dea.plot((1/eco_est01b$eff.bc)*x,z, add = T, RTS="crs",lty = "dashed", col= "blue")

# Variação da Fronteira

dea.plot((1/eco_est01b$conf.int[,1]-eco_est01b$bias)*x,z, add = T, RTS="crs", lty = "dotted", col = "red")

#dea.plot((1/evo_est01b$conf.int[,1]+evo_est01b$bias)*x,z, add = T, RTS="vrs", ORIENTATION = "out", lty =
"dashed", col = "red") <- Não precisa

#Para análise dos intervalores de Confiança exportar:

boot1 <- data.frame(evo_est01b$eff, evo_est01b$bias, evo_est01b$eff.bc, evo_est01b$conf.int)

#Exportando o Banco de Dados para Excel (Resultado 01)

library("writexl")

write_xlsx(boot1,"boot1.xlsx")

##### Verificar o Lambda (Benchmarks) e as Folgas

library("writexl")

#Selecionar abaixo o pressuposto e criar e Exportar bancos de dados:

#Para modelo CRS-Input

```

```

Res1_eci_Lambda <- data.frame(eci_est01$eff, eci_est01$slack, eci_est01$sum,
eci_est01$sx,eci_est01$sy,eci_est01$lambda)

write_xlsx(Res1_eci_Lambda,"Res1_eci_Lambda.xlsx")

#Para modelo CRS-Output

Res1_eco_Lambda <- data.frame(eco_est01$eff, eco_est01$slack, eco_est01$sum,
eco_est01$sx,eco_est01$sy,eco_est01$lambda)

write_xlsx(Res1_eco_Lambda,"Res1_eco_Lambda.xlsx")

#Para modelo VRS-Input

Res1_evi_Lambda <- data.frame(evi_est01$eff, evi_est01$slack, evi_est01$sum, evi_est01$sx,evi_est01$sy,
evi_est01$lambda)

write_xlsx(Res1_evi_Lambda,"Res1_evi_Lambda.xlsx")

#Para modelo VRS-Output

Res1_evo_Lambda <- data.frame(evo_est01$eff, evo_est01$slack, evo_est01$sum, evo_est01$sx,evo_est01$sy,
evo_est01$lambda)

write_xlsx(Res1_evo_Lambda,"Res1_evo_Lambda.xlsx")

#####ANÁLISE DA ETAPA 02#####

#Criando a matriz de PRODUTOS FINAIS.

y <- as.matrix(with(M4, Y1))

#Verificando Outliers

outliers2 <- outlier.ap (z, y, NDEL = 5, NLEN = 25, TRANSPOSE = FALSE)

outlier.ap.plot(outliers2$ratio)

outliers2$imat

#O principal outlier é a organização 13.

#Primeiramente utilizaremos o DEA tradicional - com benchmarks e folgas (slacks) - para calcular todos os índices
em todos os pressupostos:

## Retornos constantes à escala e orientação insumo (Eficiência para Retorno Constante com orientação para Input
- ECI)

eci_est02 <- dea(z,y, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)

## Retornos constantes à escala e orientação produto (Eficiência para Retorno Constante com orientação para
Output - ECO)

eco_est02 <- dea(z,y, RTS="crs", ORIENTATION = "out", SLACK=TRUE)

## Retornos variáveis à escala e orientação insumo (Eficiência para Retorno Variável com orientação para Input -
EVI)

evi_est02 <- dea(z,y, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)

## Retornos variáveis à escala e orientação produto (Eficiência para Retorno Variável com orientação para Output
- EVO)

evo_est02 <- dea(z,y, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

```



```

#Em seguida vamos rodar o DEA bootstrap para cada caso. O número recomendado de repetições é 2000 (STN,
2015)

eci_est02b <- dea.boot(z,y, NREP = 10000, RTS="crs", ORIENTATION = "in", alpha = 0.05)
eco_est02b <- dea.boot(z,y, NREP = 10000, RTS="crs", ORIENTATION = "out", alpha = 0.05)
evi_est02b <- dea.boot(z,y, NREP = 10000, RTS="vrs", ORIENTATION = "in", alpha = 0.05)
evo_est02b <- dea.boot(z,y, NREP = 10000, RTS="vrs", ORIENTATION = "out", alpha = 0.05)

#Por fim, avaliar as supereficiências para distinguir níveis de desempenho entre os eficientes.

eci_est02s <- sdea(z,y, RTS = "crs", ORIENTATION = "in")
eco_est02s <- sdea(z,y, RTS="crs", ORIENTATION = "out")
evi_est02s <- sdea(z,y, RTS="vrs", ORIENTATION = "in")
evo_est02s <- sdea(z,y, RTS="vrs", ORIENTATION = "out")

# Combinando os resultados em um banco de dados.

# CRS = Constant Returns to Scale (retornos constantes) e VRS = Variable returns to Scale (Retornos variáveis),
# in= orientação input, ou= orientação output

# Cálculo das inversas (para melhor interpretação) e cálculo das ineficiências causadas pela escala de produção
(crs/vrs).

# Do bootstrap e do sdea traremos apenas as eficiências ajustadas por essas duas técnicas.

res2 <- data.frame(crs02_in = eci_est02$eff, crs02_ou = eco_est02$eff, vrs02_in = evi_est02$eff, vrs02_ou =
evo_est02$eff, #eficiências

                 crs02_1ou = 1/eco_est02$eff, vrs02_1ou = 1/evo_est02$eff, #inversa

                 esc02_in = eci_est02$eff/evi_est02$eff, esc02_ou = eco_est02$eff/evo_est02$eff, #eficiência de
escala de in e out

                 crs02_inb = eci_est02b$eff.bc, crs02_oub = eco_est02b$eff.bc, #bootstrap
                 vrs02_inb = evi_est02b$eff.bc, vrs02_oub = evo_est02b$eff.bc, #bootstrap

                 scrs02_in = eci_est02s$eff, scrs02_ou = eco_est02s$eff, svrs02_in = evi_est02s$eff, svrs02_ou =
evo_est02s$eff) #Superficiências

#Exportando o Banco de Dados para Excel (Resultado 02)

library("writexl")

write_xlsx(res2,"Res2.xlsx")

#####Verificando suposição de retornos à escala com teste estatístico

#Para prosseguir com a análise e definição da Fronteira de eficiência é preciso definir quais pressupostos serão
utilizados. Com relação à orientação se utilizará, em razão do problema de pesquisa, a orientação output. As
Unidades de capacitação no setor público federal recebem um orçamento fixo anual para ser dispendido da melhor
forma com treinamento. Assim, o que se busca é a otimização dos outputs, sendo essa a orientação adotada.

#Com relação aos retornos, é preciso verificar se há diferenças nas distribuições probabilísticas dos scores de
eficiências entre retornos constantes e os índices de eficiência de retornos variáveis. Se houver diferença, isso
indica que os retornos são variáveis, a curva da fronteira de eficiência de produção é côncava e há ineficiência de
escala na amostra, devendo adotar-se como pressuposto a variabilidade dos retornos (VRS) - o retorno tende a cair
com o aumento da escala de produção.

```

#Teste de Kolmogorov-Smirnov - Não paramétrica para análise de igualdade entre distribuições de probabilidade contínuas de amostras. O Teste KS normalmente é utilizado para comparação de uma distribuição com uma distribuição Normal para se saber se a primeira atende a esse pressuposto. No presente caso, a comparação será feita entre distribuições de resultados de índices de eficiência, visando atestar se elas são distintas.

#Ho: Não há diferença = Não há ausência de ineficiência de escala.

(o modelo com a pressuposição de retornos constantes pode ser adotado);

#Hi: Há diferença = Há ineficiência de escala

(o modelo com a pressuposição de retornos variáveis é o mais adequado para a análise);

#O valor crítico da estatística "D" do teste KS para uma amostra de 30 casos é de 0,24 ao $p = 0,05$. Se $D > 0,24$ rejeita-se "Ho".

#Teste dos Scores Normais para input e para output:

ks.test(eqi_est02\$eff, evi_est02\$eff, alternative = "two.sided", exact = NULL)

ks.test(eco_est02\$eff, evo_est02\$eff, alternative = "two.sided", exact = NULL)

#Uma vez definida a Escala mais adequada, plotar os dados. (No caso "VRS-Output")

#####Plotando Resultados

#Isoquanta para os dois insumos

#dea.plot.isoquant(M4\$X1, M4\$X2, RTS="vrs", txt=T) Não se aplica

#Quando houver dois ou mais produtos - Função Transformação. Não se aplica

#####Plotando Resultados

#Isoquanta para os dois insumos

dea.plot.isoquant(M4\$X1, M4\$X2, RTS="vrs", txt=M4\$CASO)

#Quando houver dois ou mais produtos - Função Transformação. Não é o caso.

#Função para modalidade

Fronteira de Retornos Constantes

dea.plot.frontier(z, y, RTS="crs", txt=T)

dea.plot(z,y, RTS="crs", txt=M4\$CASO, fex=0.9, pch=17)

Fronteira de Retornos Variáveis

dea.plot(z,y, RTS="vrs", txt=M4\$CASO, fex=0.9, pch=17, cex.axis=0.9, xlab="Horas", ylab = "Participações")

Misturando as Fronteiras

dea.plot.frontier(z, y, RTS="crs", add=T)

#####Plotando intervalos de confiança oriundos do Bootstrap

#Criando um gráfico com intervalos de confiança:

1º Gráfico Original

dea.plot(Z,Y, RTS="vrs", ORIENTATION = "out", txt=T)

#Fronteira VRS

Fronteira Corrigida

```

dea.plot((1/evo_est02b$eff.bc)*z,y, add = T, RTS="vrs", lty = "dashed", col= "blue")

# Variação da Fronteira
dea.plot((1/evo_est02b$conf.int[,1]-evo_est02b$bias)*z,y, add = T, RTS="vrs", lty = "dotted", col = "red")

#Fronteira CRS
# Fronteira Corrigida
dea.plot((1/eco_est02b$eff.bc)*z,y, add = T, RTS="crs",lty = "dashed", col= "blue")

# Variação da Fronteira
dea.plot((1/eco_est02b$conf.int[,1]-eco_est02b$bias)*z,y, add = T, RTS="crs", lty = "dotted", col = "red")

#dea.plot((1/evo_est01b$conf.int[,1]+evo_est01b$bias)*x,z, add = T, RTS="vrs", ORIENTATION = "out", lty =
"dashed", col = "red") <- Não precisa

#Para análise dos intervalos de Confiança exportar:
boot2 <- data.frame(evo_est02b$eff, evo_est02b$bias, evo_est02b$eff.bc, evo_est02b$conf.int)

#Exportando o Banco de Dados para Excel (Resultado 01)
library("writexl")
write_xlsx(boot2,"boot2.xlsx")

##### Verificar o Lambda (Benchmarks) e as Folgas
library("writexl")

#Selecionar abaixo o pressuposto e criar e Exportar bancos de dados:
#Para modelo CRS-Input
Res2_eci_Lambda <- data.frame(eci_est02$eff, eci_est02$slack, eci_est02$sum,
eci_est02$sx,eci_est02$sy,eci_est02$lambda)
write_xlsx(Res2_eci_Lambda,"Res2_eci_Lambda.xlsx")

#Para modelo CRS-Output
Res2_eco_Lambda <- data.frame(eco_est02$eff, eco_est02$slack, eco_est02$sum,
eco_est02$sx,eco_est02$sy,eco_est02$lambda)
write_xlsx(Res2_eco_Lambda,"Res2_eco_Lambda.xlsx")

#Para modelo VRS-Input
Res2_evi_Lambda <- data.frame(evi_est02$eff, evi_est02$slack, evi_est02$sum, evi_est02$sx,evi_est02$sy,
evi_est02$lambda)
write_xlsx(Res2_evi_Lambda,"Res2_evi_Lambda.xlsx")

#Para modelo VRS-Output
Res2_evo_Lambda <- data.frame(evo_est02$eff, evo_est02$slack, evo_est02$sum, evo_est02$sx,evo_est02$sy,
evo_est02$lambda)
write_xlsx(Res2_evo_Lambda,"Res2_evo_Lambda.xlsx")

```

APÊNDICE H - *Script* final no software R: variáveis não-discricionárias

Testes de Análise de Fatores

#Esse script investiga as diferenças de eficiência em grupos categorizados com base na incidência de fatores interferentes. Análise de possíveis fatores setoriais e organizacionais determinantes da (in)eficiência com variáveis dicotomizadas para comparação de grupos:

#Será utilizada a proposta original de Charnes, Rhodes e Cooper (1981) de comparação não-paramétrica sugerida como a forma baseada em DEA de comparação (Boueri et al. 2015). Há outras formas paramétricas como as regressões tobit, mas devido à natureza da amostra e a necessidade de pressupostos de análise optou-se pelo modelo original de CCR.

Verificar o impacto dos fatores por meio do DEA

#Em situações nas quais a amostra pode ser dividida em subamostras de acordo com alguma variável categórica (e.g. administração pública e privada), Charnes, Cooper e Rhodes (1981) propõem o seguinte procedimento para análise comparativa de eficiência segundo essa variável, em particular:

1. Divida a amostra nas diversas subamostras compatíveis e rode um DEA para cada subamostra;

2. Projete cada ponto na fronteira obtida para a respectiva subamostra;

 #(i.e. simule todas DMUs como eficientes corrigindo sua ineficiência por meio da multiplicação de seus insumos/produtos pelo índice de eficiência correspondente)

3. Rode mais um DEA, desta vez com as projeções (misturadas), e calcule as médias de eficiências de cada categoria.

#A lógica é que se avaliará se um dos grupos, assumindo-se a máxima eficiência possível de todos os casos, apresentará uma eficiência média maior que o outro grupo. Isso demonstra que o grupo ineficiente poderia incrementar sua eficiência no valor da respectiva diferença entre as médias, caso tivesse a mesma condição ambiental do outro grupo.

Script Geral para Análise de Fatores

#Carregando a Planilha - O modelo selecionado é o QUARTO modelo testado e melhor ajustado à amostra, pressupostos metodológicos e objetivos de pesquisa.

```
library(readxl)
```

```
M4 <- read_excel("M4.xlsx")
```

```
#Iremos utilizar o pacote Benchmarking
```

```
library(Benchmarking)
```

```
#Verificar os nomes das colunas
```

```
colnames(M4)
```

```
#Devolva o nome original do fator "F"
```

```
#names(M4)[names(M4)=="F"] <- "FPAC"
```

```
#Verificar novamente os nomes das colunas
```

```
colnames(M4)
```

```
#Modificar o nome do próximo fator a ser analisado na planilha para "F"
```

```

names(M4)[names(M4)== "FPORT"] <- "F"
#Verificar os nomes das colunas antes da análise
colnames(M4)
#Criando subgrupos de insumos "x01" e produtos "z01" e "y01" dos casos com valor "1" para o fator.
x01 <- as.matrix(subset(M4, F==1,select = c(X1,X2)))
z01 <- as.matrix(subset(M4, F==1,select = Z))
y01 <- as.matrix(subset(M4, F==1,select = Y1))
#Criando subgrupos de insumos "x01" e produtos "z01" e "y01" dos casos com valor "0" para o fator.
x00 <- as.matrix(subset(M4, F==0,select = c(X1,X2)))
#x002 <- as.matrix(subset(M4, FCOMP==0,select = (X2)))
z00 <- as.matrix(subset(M4, F==0,select = Z))
y00 <- as.matrix(subset(M4, F==0,select = Y1))
#####Rodando os DEAs para o grupo 01 "com" o fator.
#####ESTÁGIO 01
## Retornos constantes à escala e orientação insumo
eci_est01G1 <- dea(x01,z01, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)
## Retornos constantes à escala e orientação produto
eco_est01G1 <- dea(x01,z01, RTS="crs", ORIENTATION = "out", SLACK=TRUE)
## Retornos variáveis à escala e orientação insumo
evi_est01G1 <- dea(x01,z01, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)
## Retornos variáveis à escala e orientação produto
evo_est01G1 <- dea(x01,z01, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
#Modelo utilizado: VRS orientado para output (evo)
res1G1 <- data.frame(x01, z01, x01a = x01*evo_est01G1$sx, z01a = z01*evo_est01G1$eff, # Corrigindo e
criando a versão ajustada para eficiência máxima "x01a" e "z01a".
                crs01_in = eci_est01G1$eff, crs01_ou = eco_est01G1$eff, vrs01_in = evi_est01G1$eff, vrs01_ou =
evo_est01G1$eff, #eficiências
                esc01_in = eci_est01G1$eff/evi_est01G1$eff, esc01_ou = eco_est01G1$eff/evo_est01G1$eff, #escala
de in e out
                evo_est01G1$slack, evo_est01G1$sum, evo_est01G1$sx,evo_est01G1$sy,evo_est01G1$lambda)
#Lambda do EVO
#checando a variável nova
names(res1G1)
#Criando insumos e produtos ajustados para a eficiência:
x01a <- as.matrix(with(res1G1, cbind(x01a.X1,x01a.X2)))
z01a <- as.matrix(with(res1G1, Z.1))

```

```

#Conferindo as eficiências originais:
evo_est01G1 <- dea(x01a,z01, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
evo_est01G1$eff

#Contra as eficiências ajustadas:
evo_est01G1a <- dea(x01a,z01a, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
evo_est01G1a$eff # < Todas têm que ser = 1

#Exportando o Banco de Dados para Excel.
library("writexl")
write_xlsx(res1G1,"Res1G1.xlsx")

#####Rodando os DEAs para o grupo "0" sem o fator
## Retornos constantes à escala e orientação insumo
eci_est01G0 <- dea(x00,z00, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)
## Retornos constantes à escala e orientação produto
eco_est01G0 <- dea(x00,z00, RTS="crs", ORIENTATION = "out", SLACK=TRUE)
## Retornos variáveis à escala e orientação insumo
evi_est01G0 <- dea(x00,z00, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)
## Retornos variáveis à escala e orientação produto
evo_est01G0 <- dea(x00,z00, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

res1G0 <- data.frame(x00, z00, x00a = x00-evo_est01G0$sx, z00a = z00*evo_est01G0$eff,
                    crs01_in = eci_est01G0$eff, crs01_ou = eco_est01G0$eff, vrs01_in = evi_est01G0$eff, vrs01_ou =
                    evo_est01G0$eff, #eficiências
                    esc01_in = eci_est01G0$eff/evi_est01G0$eff, esc01_ou = eco_est01G0$eff/evo_est01G0$eff,
                    #escala de in e out
                    evo_est01G0$slack, evo_est01G0$sum, evo_est01G0$sx,evo_est01G0$sy,evo_est01G0$lambda)
#Lambda do EVO

#checando a variável nova
names(res1G0)

#Criando insumos e produtos ajustados:
x00a <- as.matrix(with(res1G0, cbind(x00a.X1,x00a.X2)))
z00a <- as.matrix(with(res1G0, Z.1))

#Conferindo as eficiências originais:
evo_est01G0$eff

#Contra as eficiências ajustadas:
evo_est01G0a <- dea(x00a,z00a, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
evo_est01G0a$eff #< Todas = 1 Ok

```

```

#Exportando o Banco de Dados para Excel
library("writexl")
write_xlsx(res1G0,"Res1G0.xlsx")
##### Comparar os plots
dea.plot(x01,z01, RTS="Vrs", ORIENTATION = "out",col="blue", pch=16, txt = M4$CASO)
dea.plot(x01,z01, RTS="Vrs", ORIENTATION = "out",col="blue", pch=16, txt = M4$CASO)
dea.plot(x00,z00, RTS="Vrs", add = TRUE, ORIENTATION = "out",col="red", pch=1)
X00.x2 <- res1G0$X2
points(X00.x2, z00, col = "red", pch=1)
#text(x, z, 1:dim(x)[1], adj=-1)
dea.plot(x01a,z01a, RTS="Vrs", ORIENTATION = "out",col="blue", pch=16, txt = M4$CASO)
dea.plot(x00a,z00a, RTS="Vrs", add = TRUE, ORIENTATION = "out",col="red", pch=1)
##### Rodar o DEA com todos junto com eficiência máxima
#Verificando nome das variáveis X e Z ajustadas
colnames(res1G1)
colnames(res1G0)
#Criando dados de insumo-produto ajustados para eficiência máxima e identificando o F.
res1G1a <- data.frame(X1a = res1G1$x01a.X1, X2a = res1G1$x01a.X2, Za = res1G1$Z.1, "F" = 1)
res1G0a <- data.frame(X1a = res1G0$x00a.X1, X2a = res1G0$x00a.X2, Za = res1G0$Z.1, "F" = 0)
#Juntando ambos em um só data set
res1Ga <- data.frame(rbind(res1G1a,res1G0a))
# separando insumos e produtos intermediário Z
xGa <- as.matrix(with(res1Ga, cbind(X1a, X2a)))
zGa <- as.matrix(with(res1Ga, Za))
## Rodando o DEA integrado com Retornos variáveis à escala e orientação produto
evo_est01Ga <- dea(xGa,zGa, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
#Res1Ga <- data.frame(res1Ga$X1a, res1gpca$X2a, res1gpca$Za, res1gpca$G,
#
#                               evo_est01GPC1$eff, evo_est01GPC1$slack, evo_est01GPC1$sum,
evo_est01GPC1$sx,evo_est01GPC1$sy, evo_est01GPC1$lambda)
#Exportando a comparação entre os dois grupos
#write_xlsx(Res1GPC1,"Res1GPC1.xlsx")
#####incorporando os testes ao BANCO
res1Ga <- cbind(res1Ga,Eff = evo_est01Ga$eff)
#####Separando as distribuições de eficiência para comparação)
Eff01 <- as.matrix(subset(res1Ga, F==1,select = "Eff"))

```

```

Eff00 <- as.matrix(subset(res1Ga, F==0,select = "Eff"))
Eff01inv <- 1/Eff01
Eff00inv <- 1/Eff00
# boxplot das duas distribuições - Colocar no xlab qual o nome do fator comparado.
boxplot((1/res1Ga$Eff)~res1Ga$F, las=1, ylab="Eficiência",
        xlab="ANT",main="Eficiência comparada",
        ylim = c(0.2, 1.1), yaxs = "i")
KS1 <- ks.test(Eff01, Eff00, alternative = "two.sided", exact = NULL)
#Ho: Não tem diferença
#Hi: Tem diferença
# Calculando a diferença entre as médias de eficiências das amostras que serve de base para o teste de hipótese
mean(Eff01inv)
mean(Eff00inv)
mean(1/res1Ga$Eff[res1Ga$F==1])
mean(1/res1Ga$Eff[res1Ga$F==0])
# Outra forma:
with(res1Ga, tapply((1/Eff), F, mean))
# Calculando a diferença entre essas médias
Dif.Media <- (mean(Eff01inv) - mean(Eff00inv)) #diferença de médias
Dif.Media
# Outra forma
-diff( with(res1Ga, tapply(1/Eff, F, mean)) )
# Diferença das Medianas
median(1/res1Ga$Eff[res1Ga$F==1])
median(1/res1Ga$Eff[res1Ga$F==0])
# Outra forma
with(res1Ga, tapply(1/Eff, F, median))
# lets calculate the diff in medians: (casein - meatmeal)
Dif.Mediana <- (median(Eff01inv) - median(Eff00inv))
Dif.Mediana
# and, a fancier way to do that... (- to have it be casein-meatmeal)
-diff( with(res1Ga, tapply(1/Eff, F, median)) )
### INTERVALO DE CONFIANÇA COM BOOTSTRAP
# Criar uma seed para o bootstrap ser consistente e reproduzível

```



```

setseed(13579)

n.1 <- NROW(z01) # n de observações para o grupo 1
n.0 <- NROW(z00) # n de observações para o grupo 0
B <- 100000 # número de reamostragens bootstrap
#criando a função para rodar o Bootstrap
Boot.01 <- matrix( sample(Eff01inv, size= B*n.1,
                        replace=TRUE), ncol=B, nrow=n.1)
Boot.00 <- matrix( sample(Eff00inv, size= B*n.0,
                        replace=TRUE), nrow=n.0, ncol=B)

# Verificar se dimensões estão corretas
dim(Boot.01); dim(Boot.00)

# Verificar se os resultados foram gerados
Boot.01[1:2,1:2]
Boot.00[1:2,1:2]

# Calcular as diferenças entre as médias dos grupos.
Boot.Dif.Media <- colMeans(Boot.01) - colMeans(Boot.00)

# Verifiar o resultados
length(Boot.Dif.Media)
Boot.Dif.Media[1:10]

# Calcular as diferenças entre as medianas dos grupos
Boot.Dif.Mediana <- apply(Boot.01, MARGIN=2, FUN=median) -
  apply(Boot.00, MARGIN=2, FUN=median)

# Verificar resultados
length(Boot.Dif.Mediana)
Boot.Dif.Mediana[1:10]

#### CRIANDO INTERVALOS DE CONFIANÇA (A 95%)
# MÉTODO DE PERCENTIL:
# Para diferença das Médias
quantile(Boot.Dif.Media, prob=0.025)
quantile(Boot.Dif.Media, prob=0.975)
# Para diferença das Medianas
quantile(Boot.Dif.Mediana, prob=0.025)
quantile(Boot.Dif.Mediana, prob=0.975)

```

```

#####Escrever a diferença

#O efeito do fator pode ser em média de:

Dif.Media

#Variando de:

Min <- quantile(Boot.Dif.Media, prob=0.025)

#Até:

Max <- quantile(Boot.Dif.Media, prob=0.975)

#####Escrever a diferença

F1 <- data.frame(Q1 = n.1, Q0 = n.0, G1 = mean(Eff01inv), G0 = mean(Eff00inv), Dif.Media, Min, Max,
KS1$p.value)

write_xlsx(F1,"F1.xlsx")

#####ESTÁGIO 02#####

## Retornos constantes à escala e orientação insumo

eci_est02G1 <- dea(z01,y01, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)

## Retornos constantes à escala e orientação produto

eco_est02G1 <- dea(z01,y01, RTS="crs", ORIENTATION = "out", SLACK=TRUE)

## Retornos variáveis à escala e orientação insumo

evi_est02G1 <- dea(z01,y01, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)

## Retornos variáveis à escala e orientação produto

evo_est02G1 <- dea(z01,y01, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

#Modelo utilizado: VRS orientado para output (evo)

res2G1 <- data.frame(z01, y01, z01a = z01-evo_est02G1$sx, y01a = y01*evo_est02G1$eff, # Corrigindo e
criando a versão ajustada para eficiência máxima "x01a" e "z01a".

          crs01_in = eci_est02G1$eff, crs01_ou = eco_est02G1$eff, vrs01_in = evi_est02G1$eff, vrs01_ou =
evo_est02G1$eff, #eficiências

          esc01_in = eci_est02G1$eff/evi_est02G1$eff, esc01_ou = eco_est02G1$eff/evo_est02G1$eff,
#escala de in e out

          evo_est02G1$slack, evo_est02G1$sum, evo_est02G1$sx,evo_est02G1$sy,evo_est02G1$lambda)
#Lambda do EVO

#checando a variável nova

names(res2G1)

#Criando insumos e produtos ajustados para a eficiência:

z01a <- as.matrix(with(res2G1, Z.1))

y01a <- as.matrix(with(res2G1, Y1.1))

#Conferindo as efficiências originais:

evo_est02G1 <- dea(z01,y01, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

```

```

evo_est02G1$eff

#Contra as eficiências ajustadas:
evo_est02G1a <- dea(z01a,y01a, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
evo_est02G1a$eff # < Todas têm que ser = 1
#Exportando o Banco de Dados para Excel.
library("writexl")
write_xlsx(res2G1,"Res2G1.xlsx")
#####Rodando os DEAs para o grupo "0" sem o fator
## Retornos constantes à escala e orientação insumo
eci_est02G0 <- dea(z00,y00, RTS = "crs", ORIENTATION = "in", SLACK=TRUE)
## Retornos constantes à escala e orientação produto
eco_est02G0 <- dea(z00,y00, RTS="crs", ORIENTATION = "out", SLACK=TRUE)
## Retornos variáveis à escala e orientação insumo
evi_est02G0 <- dea(z00,y00, RTS="vrs", ORIENTATION = "in", SLACK=TRUE)
## Retornos variáveis à escala e orientação produto
evo_est02G0 <- dea(z00,y00, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
res2G0 <- data.frame(z00, y00, z00a = z00-evo_est02G0$sx, y00a = y00*evo_est02G0$eff,
                    crs02_in = eci_est02G0$eff, crs02_ou = eco_est02G0$eff, vrs02_in = evi_est02G0$eff, vrs02_ou =
evo_est02G0$eff, #eficiências
                    esc02_in = eci_est02G0$eff/evi_est02G0$eff, esc02_ou = eco_est02G0$eff/evo_est02G0$eff,
#escala de in e out
                    evo_est02G0$slack, evo_est02G0$sum, evo_est02G0$sx,evo_est02G0$sy,evo_est02G0$lambda)
#Lambda do EVO
#checando a variável nova
names(res2G0)
#Criando insumos e produtos ajustados:
z00a <- as.matrix(with(res2G0, Z.1))
y00a <- as.matrix(with(res2G0, Y1.1))
#Conferindo as eficiências originais:
evo_est02G0$eff
#Contra as eficiências ajustadas:
evo_est02G0a <- dea(z00a,y00a, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)
evo_est02G0a$eff #< Todas = 1 Ok
#Exportando o Banco de Dados para Excel
library("writexl")

```

```

write_xlsx(res2G0,"Res2G0.xlsx")

##### Comparar os plots
dea.plot(z01,y01, RTS="Vrs", ORIENTATION = "out",col="blue", pch=16, txt = M4$CASO)
dea.plot(z01,y01, RTS="Vrs", ORIENTATION = "out",col="blue", pch=16, txt = M4$CASO)
dea.plot(z00,y00, RTS="Vrs", add = TRUE, ORIENTATION = "out",col="red", pch=1)
points(z00, y00, col = "red", pch=1)
#text(x, z, 1:dim(x)[1], adj=-1)
dea.plot(z01a,y01a, RTS="Vrs", col="blue", pch=16, txt = M4$CASO)
dea.plot(z00a,y00a, RTS="Vrs", add = TRUE, col="red", pch=1)

##### Rodar o DEA com todos junto com eficiência máxima
#Verificando nome das variáveis X e Z ajustadas
colnames(res2G1)
colnames(res2G0)

#Criando dados de insumo-produto ajustados para eficiência máxima e identificando o F.
res2G1a <- data.frame(za = res2G1$Z.1, ya = res2G1$Y1.1, "F" = 1)
res2G0a <- data.frame(za = res2G0$Z.1, ya = res2G0$Y1.1, "F" = 0)

#Juntando ambos em um só data set
res2Ga <- data.frame(rbind(res2G1a,res2G0a))

# separando insumos e produtos intermediário Z
zGa <- as.matrix(with(res2Ga, za))
yGa <- as.matrix(with(res2Ga, ya))

## Rodando o DEA integrado com Retornos variáveis à escala e orientação produto
evo_est02Ga <- dea(zGa,yGa, RTS="vrs", ORIENTATION = "out", SLACK=TRUE)

#Res1Ga <- data.frame(res1Ga$X1a, res1gpca$X2a, res1gpca$Za, res1gpca$G,
#
#               evo_est01GPC1$eff, evo_est01GPC1$slack, evo_est01GPC1$sum,
evo_est01GPC1$ss,evo_est01GPC1$sy, evo_est01GPC1$lambda)

#Exportando a comparação entre os dois grupos
#write_xlsx(Res1GPC1,"Res1GPC1.xlsx")

#####incorporando os testes ao BANCO
res2Ga <- cbind(res2Ga,Eff = evo_est02Ga$eff)

#####Separando as distribuições de eficiência para comparação)
Eff201 <- as.matrix(subset(res2Ga, F==1,select = "Eff"))
Eff200 <- as.matrix(subset(res2Ga, F==0,select = "Eff"))
Eff201inv <- 1/Eff201
Eff200inv <- 1/Eff200

```

```

# boxplot das duas distribuições - Colocar no xlab qual o nome do fator comparado.
boxplot((1/res2Ga$Eff)~res2Ga$F, las=1, ylab="Eficiência",
        xlab="ANT",main="Eficiência comparada",
        ylim = c(0.2, 1.1), yaxs = "i")
KS2 <- ks.test(Eff201, Eff200, alternative = "two.sided", exact = NULL)
#Ho: Não tem diferença
#Hi: Tem diferença
# Calculando a diferença entre as médias de eficiências das amostras que serve de base para o teste de hipótese
mean(Eff201inv)
mean(Eff200inv)
mean(1/res2Ga$Eff[res1Ga$F==1])
mean(1/res2Ga$Eff[res1Ga$F==0])
# Outra forma:
with(res2Ga, tapply((1/Eff), F, mean))
# Calculando a diferença entre essas médias
Dif2.Media <- (mean(Eff201inv) - mean(Eff200inv)) #diferença de médias
Dif2.Media
# Outra forma
-diff( with(res2Ga, tapply(1/Eff, F, mean)) )
# Diferença das Medianas
median(1/res2Ga$Eff[res2Ga$F==1])
median(1/res2Ga$Eff[res2Ga$F==0])
# Outra forma
with(res2Ga, tapply(1/Eff, F, median))
# lets calculate the diff in medians: (casein - meatmeal)
Dif2.Mediana <- (median(Eff201inv) - median(Eff200inv))
Dif2.Mediana
# and, a fancier way to do that... (- to have it be casein-meatmeal)
-diff( with(res2Ga, tapply(1/Eff, F, median)) )
### INTERVALO DE CONFIANÇA COM BOOTSTRAP
# Criar uma seed para o bootstrap ser consistente e reproduzível
setseed(13580)
n.1 <- NROW(z01) # n de observações para o grupo 1
n.0 <- NROW(z00) # n de observações para o grupo 0

```

```

B <- 100000 # número de reamostragens bootstrap
#criando a função para rodar o Bootstrap
Boot2.01 <- matrix( sample(Eff201inv, size= B*n.1,
                          replace=TRUE), ncol=B, nrow=n.1)
Boot2.00 <- matrix( sample(Eff200inv, size= B*n.0,
                          replace=TRUE), nrow=n.0, ncol=B)
# Verificar se dimensões estão corretas
dim(Boot2.01); dim(Boot2.00)
# Verificar se os resultados foram gerados
Boot2.01[1:2,1:2]
Boot2.00[1:2,1:2]
# Calcular as diferenças entre as médias dos grupos.
Boot2.Dif.Media <- colMeans(Boot2.01) - colMeans(Boot2.00)
# Verificar o resultados
length(Boot2.Dif.Media)
Boot2.Dif.Media[1:10]
# Calcular as diferenças entre as medianas dos grupos
Boot2.Dif.Mediana <- apply(Boot2.01, MARGIN=2, FUN=median) -
  apply(Boot2.00, MARGIN=2, FUN=median)
# Verificar resultados
length(Boot2.Dif.Mediana)
Boot2.Dif.Mediana[1:10]
#### CRIANDO INTERVALOS DE CONFIANÇA (A 95%)
# MÉTODO DE PERCENTIL:
# Para diferença das Médias
quantile(Boot2.Dif.Media, prob=0.025)
quantile(Boot2.Dif.Media, prob=0.975)
# Para diferença das Medianas
quantile(Boot2.Dif.Mediana, prob=0.025)
quantile(Boot2.Dif.Mediana, prob=0.975)
#####Escrever a diferença
#O efeito do fator pode ser em média de:
Dif2.Media
#Variando de:

```

```
Min2 <- quantile(Boot2.Dif.Media, prob=0.025)
```

```
Min2
```

```
#Até:
```

```
Max2 <- quantile(Boot2.Dif.Media, prob=0.975)
```

```
Max2
```

```
#####Escrever a diferença
```

```
F2 <- data.frame(Q1 = n.1, Q0 = n.0, G1 = mean(Eff201inv), G0 = mean(Eff200inv), Dif2.Media, Min2, Max2,  
KS2$p.value)
```

```
write_xlsx(F2,"F2.xlsx")
```