

**A Avaliação baseada na Teoria na Avaliação de Impacto em
Programas Complexos: propostas de inovação na aplicação da
teoria da mudança e no teste dos mecanismos de causalidade”**

Equipa: Paulo Pedroso (Coordenador), Paulo Feliciano, Teresa Evaristo,

Ana Sofia Amândio, Filipa Seiceira, Maria Álvares, Otávio Moreira Junior

Relatório Final

AVISO PARA APRESENTAÇÃO DE CANDIDATURAS N.º 06/POAT/2021

Agência para o Desenvolvimento e Coesão

Programa Operacional de Assistência Técnica Portugal 2020

Índice

INTRODUÇÃO	4
1. A AVALIAÇÃO BASEADA NA TEORIA: PRINCIPAIS ABORDAGENS E CONTRIBUTOS	5
<i>Caixa negra</i>	<i>8</i>
<i>Análise generativa</i>	<i>9</i>
<i>Análise de contribuição.....</i>	<i>10</i>
<i>FEEI</i>	<i>10</i>
2. A TDM NA PERSPETIVA DA COMUNIDADE DE PRÁTICAS.....	12
<i>Discussão dos resultados</i>	<i>13</i>
<i>Programação</i>	<i>14</i>
<i>Operacionalização</i>	<i>14</i>
<i>Dados empíricos.....</i>	<i>15</i>
3. A CARACTERIZAÇÃO DAS INTERVENÇÕES COMPLEXAS E OS DESAFIOS QUE COLOCA À AVALIAÇÃO	18
3.1 A AVALIAÇÃO REALISTA.....	27
3.2 AS TDM “ENCAIXADAS”	29
<i>Notas conclusivas.....</i>	<i>34</i>
4. A INVESTIGAÇÃO DA AÇÃO SOCIAL E O CONCEITO DE MECANISMO	35
<i>Process Tracing</i>	<i>39</i>
<i>Aplicações de Análise de Configurações Causais</i>	<i>43</i>
5. OPERACIONALIZAÇÃO DA ANÁLISE DE CONFIGURAÇÕES CAUSAIS (PROCESS TRACING).....	46
4.1 TEORIA DA MUDANÇA E ANÁLISE DE CONFIGURAÇÕES CAUSAIS — O CASO PRÁTICO	52
6. SÍNTESE FINAL	68
REFERÊNCIAS BIBLIOGRÁFICAS	71
ANEXO	76

Índice de Tabelas

Tabela 1. Os Atributos das Intervenções Complexas e os Instrumentos de Política dos FEEI	21
Tabela 2. A TdM e a Avaliação Realista na Avaliação de Intervenções Complexas	28
Tabela 3. Teoria da Mudança à escala da TI, abordagem generativa, realista.	32
Tabela 4. TdM aplicada aos FEEI – versão simplificada	33
Tabela 5. Exemplo de TdMs encaixadas dos FEEI	34
Tabela 6. Tipos de Mecanismos	37
Tabela 7. Quadrantes de robustez da hipótese causal	42
Tabela 8. Operacionalização do mecanismo causal	48
Tabela 9. Proposta de sistematização de aplicação dos testes.....	49
Tabela 10. Operacionalização dos mecanismos causais da TO 1.01 – Estágios para adultos e 1.02 – Apoios à contratação de adultos.....	59
Tabela 11. Análise do peso inferencial das evidências - aplicação dos testes.....	63

Índice de Figuras

Figura 1.Dimensões da complexidade empírica da TdM aplicada ao PT2020	13
Figura 2. TdM – modelo das “teorias encaixadas”	30
Figura 4. Representação esquemática de um mecanismo causal	47
Figura 5. Sistematização dos passos de ACC.....	50
Figura 6. Esquema da Teoria da Mudança das TO 1.01 – Estágios para adultos e 1.02 – Apoios à contratação de adultos	54
Figura 7. Representação esquemática do mecanismo 1.....	56
Figura 8. Representação esquemática do mecanismo 2.....	57
Figura 9. Desagregação do mecanismo 3.....	58
Figura 10. Representação esquemática do mecanismo 4.....	58
Figura 11. Representação esquemática do mecanismo 5.....	59
Figura 12 – Teoria da Mudança da estratégia de avaliação do Portugal 2020	76
Figura 13 – Tipos de avaliações e critérios de avaliação.....	77

Introdução

O estudo “A Avaliação baseada na Teoria na Avaliação de Impacto em Programas Complexos: propostas de inovação na aplicação da teoria da mudança e no teste dos mecanismos de causalidade” compreende como objeto analisar e desenvolver abordagens metodológicas de avaliação referidas à abordagem baseada na teoria no âmbito da avaliação da eficácia e impacto de intervenções complexas no contexto da implementação dos Fundos Estruturais. Em particular, o estudo centra-se no teste dos mecanismos de causalidade a partir do rastreamento de processos (*Process Tracing*). Como adiante se explicita de forma aprofundada, entende-se por intervenções complexas as que combinam uma grande diversidade de políticas e instrumentos, envolvem diferentes níveis de implementação e visam uma grande diversidade de impactos no contexto de estratégias integradas de desenvolvimento.

O presente documento constitui o Relatório Final do estudo, sistematizando as principais análises elaboradas ao longo do trabalho e as conclusões que delas decorrem.

O primeiro capítulo desenvolve uma breve sistematização dos princípios teóricos e abordagens inerentes à avaliação baseada na teoria, em particular a proposta conceptual inerente à Teoria da Mudança (TdM). Nesta medida, constitui um referente teórico de partida para o subsequente desenvolvimento do trabalho.

O segundo capítulo reúne e analisa informação qualitativa sobre o modo como a implementação da Teoria da Mudança na prática de avaliação no âmbito dos Fundos Europeus Estruturais e de Investimento (FEEI) é vista pelos seus principais stakeholders, sobretudo no contexto da comunidade de práticas. Esta leitura permite refletir sobre as oportunidades, problemas e desafios suscitados pela adoção da TdM nos referenciais metodológicos da avaliação dos FEEI.

No terceiro capítulo problematiza-se o conceito de intervenções complexas, identificando os principais atributos que lhes conferem complexidade e o modo como essa condição interage com as abordagens metodológicas que suportam a avaliação. Situam-se neste passo as estratégias de avaliação que de forma mais explícita procuram considerar o problema da complexidade na avaliação.

O quarto capítulo elabora sobre o modo de aplicação da metodologia de *process tracing* na avaliação e o quinto ensaia a sua utilização a um caso concreto – o da avaliação de duas medidas de política ativa de emprego.

A sistematização das principais conclusões retiradas do estudo desenvolvido encerra o relatório.

Em anexo ao relatório junta-se uma apresentação onde se detalha a abordagem proposta pela metodologia de “process tracing” e que constituiu um recurso de apoio a estratégias de capacitação.

1. A Avaliação Baseada na Teoria: principais abordagens e contributos

As avaliações baseadas na teoria (AbT) visam tornar os pressupostos políticos explícitos e apresentam-se como uma nova vaga por relação às avaliações orientadas para o método (Stame, 2004).

A AbT foi cunhada por Tyler na década de 1930, remetendo para a formulação e teste da teoria do programa para fins de avaliação (Donaldson, 2007). Esta ideia reapareceu durante as décadas seguintes (Weiss, 1972; Bickman, 1987; Chen 1980; Chen & Rossi, 1980), mas o período de maior desenvolvimento dá-se em 1990, a partir dos trabalhos seminais de Chen (1990), Weiss (1997) e Pawson & Tilley (1997). Desde então, as AbT têm vindo a tornar-se cada vez mais valorizadas pelos decisores políticos para explicar por que razão uma intervenção funciona – ou não – num determinado contexto (Riché 2013). A partir destes contributos e a reflexão tem-se densificado com autores como Donaldson (2007), Rogers, (2008); Stame (2004), entre muitos outros. A avaliação baseada na teoria é descrita atualmente como "orientada pela teoria" ou pelo menos "informada pela teoria" (Stern, 2020).

A ABT tem sido cada vez mais defendida por numerosos teóricos de avaliação, profissionais e outras entidades como um dos métodos preferenciais para a prática da avaliação. Inúmeras instituições de relevo têm atribuído centralidade a esta abordagem metodológica, entre as quais a Comissão Europeia (CE) que introduziu a AbT nas orientações técnicas aplicáveis ao período de programação 2014-2020¹. Na perspetiva descrita pela CE (Evalsed, 2013), as teorias (ideias tácitas) expressam a lógica de intervenção de uma política. As ações políticas, através da atribuição (despesa) de certos recursos financeiros (*inputs*) visam produzir resultados planeados através dos quais se espera obter bem-estar e progresso. Os resultados dependem da eficácia da política e de outros fatores, como o contexto. O elemento essencial da eficácia da política nesta perspetiva, contudo, não é a cadeia input-output-resultado, o modelo lógico ou as equações estatísticas, mas antes os “mecanismos” que fazem funcionar a intervenção e que dizem respeito a crenças, desejos, cognições e outros processos que subjazem à tomada de decisão e que influenciam as escolhas e ações comportamentais.

A CE introduz a teoria na avaliação da Política de Coesão, em contraponto de abordagens antipositivistas que desvalorizam a interpretação. Reconhece, neste sentido, as perspetivas realistas e as perspetivas construtivistas, nas suas diferentes formas, sublinhando a necessidade da teoria prática: “Com a ajuda da teoria, a avaliação é capaz de analisar as intenções dos programas e identificar lógicas de intervenção; compreender os processos de implementação e mudança; e explicar os efeitos dos programas em diferentes áreas políticas.” (Evalsed, 2013: pp.

¹ Apesar de em Portugal a ABT / TDM ter sido operacionalizada anteriormente, ainda no período do QREN por Mamede & Pereira, 2018.

20, 21). Sem demarcar um posicionamento crítico, a CE especifica ainda que, na prática, as perspetivas construtivistas e realistas confundem-se, não sendo modelos estanques.

Deste modo, a CE refere-se ao termo de “avaliação realista” citando Pawson (2002) e Pawson e Tilley (1997). Esta abordagem metodológica sublinha a importância das configurações CMR (Contexto, Mecanismo, Resultados) para a análise e avaliação de políticas e programas. Distingue-se das abordagens funcionalistas por considerar central a reconstrução das teorias implícitas que se materializam nas dinâmicas de implementação. A avaliação realista visa tornar explícitas as teorias dentro de um programa, desenvolvendo hipóteses claras sobre como, e para quem, os programas podem “funcionar”. A implementação do programa, e a avaliação do mesmo, testa essas hipóteses. Tal implica recolher dados, quantitativos e qualitativos, sobre os resultados do programa, ou sobre perceções e representações sociais nos processos de implementação do programa; bem como outros aspetos específicos do contexto do programa que possam ter impacto nos resultados, e sobre os mecanismos específicos que possam criar mudanças.

A CE refere-se igualmente à perspetiva construtivista, segundo a qual as construções dos vários agentes estão dependentes dos contextos sociais, culturais, psicológicos e físicos em que emergem. A avaliação construtivista, ou abordagem avaliativa de quarta geração (Guba & Lincoln, 1989) defende que o produto da avaliação não é um conjunto de conclusões, recomendações ou julgamentos de valor, mas sim uma agenda de negociação. A ideia de que os problemas têm causas que podem ser abordadas é rejeitada pelos construtivistas. Apesar da inexistência de um posicionamento crítico, a preferência pela perspetiva realista em detrimento da perspetiva construtivista é feita pela enunciação operativa da abordagem eleita.

No âmbito das orientações técnicas sobre a AbT, a CE considera várias ferramentas e técnicas (modelos lógicos, quadros lógicos, teoria de programas, teoria da mudança) para a reconstrução de lógicas de intervenção do programa e das cadeias de implementação. Na literatura, as definições, delimitações e diferenças entre estas ferramentas nem sempre são claras. Como constata Coryn et al (2011) existe uma compreensão conceptual e operacional em grande medida elusiva. Os termos utilizados para descrever estes modelos variam muito, e incluem teoria de programas, modelo lógico, teoria da mudança, cadeia de resultados, caminho dos resultados, teoria da ação, teoria da implementação, sem qualquer acordo geral sobre os termos ou o significado. Reflexões sobre gama de termos utilizados para os diferentes modelos podem ser encontradas em Patton (2008) e Funnell & Rogers (2011).

A CE recomenda a teoria da mudança apenas como uma das opções possíveis no âmbito da ABT. Focamos aqui a TdM na medida em que a teoria da mudança é o elemento-chave das abordagens à avaliação baseadas na teoria (AbT) (Brousselle & Buregeya 2018) e desenvolveu-se por relação à avaliação de programas complexos (Kubisch et al., 1995; Stame, 2004).

Tal como no caso da ABT, apesar de haver um consenso sobre o quadro geral das teorias da mudança – que descrevem como as intervenções devem funcionar – há uma proliferação de interpretações difusas sobre o que na prática uma teoria da mudança implica, como se desenvolve, e como deve ser representada (James, 2011; Vogel, 2012; Stein & Valters, 2012).

Procurando explicitar conceitos, Weiss propõe que a TdM resulta da soma da Teoria do Programa e da Teoria da Execução. A Teoria do programa foca os processos que medeiam a

execução do programa e o alcance das metas². A Teoria de execução foca a boa execução do programa. Deste modo a TdM recobre a avaliação de todas as etapas e de todas as consequências esperadas e não esperadas (Weiss, 1972; 1999).

Os guias de orientação da CE elegem Weiss (1995) como a autora por excelência da TdM (Leeuw 2012). Para Weiss a TdM surge precisamente do confronto com a complexidade de programas e políticas e deve ser aplicada a iniciativas complexas. Weiss definiu a teoria da mudança como uma forma de descrever o conjunto de pressupostos que explicam os pequenos passos que conduzem ao objetivo de longo prazo e às ligações entre as atividades políticas ou programáticas, e os resultados que ocorrem nas diferentes etapas do processo. Os stakeholders de iniciativas complexas estão, regra geral, pouco cientes do desenrolar do processo de mudança e, por isso, descumam as mudanças de curto e médio prazo que precisam de ocorrer para que um objetivo de longo prazo seja alcançado.

Weiss desafiou os criadores de iniciativas complexas – como é o caso dos programas da UE – a especificar as TdM que orientam o seu trabalho, para melhorar as políticas e reforçar a capacidade de serem reconhecidos pelos resultados previstos na teoria. A identificação do objeto de avaliação e dos critérios de avaliação deve assentar, segundo Weiss, em objetivos claros, numa lógica coerente de intervenção, numa teoria de mudança, e numa conceção precisa do programa. Devido à existência de múltiplos objetivos e programas complexos a TdM torna-se essencial. Idealmente, as preocupações de avaliação devem ser tidas em conta na fase de formulação do programa, prevenindo problemas e objetivos conflitantes. E a conceptualização dos resultados esperados em termos operacionais e mensuráveis (sistema de monitorização e indicadores) permite também ajudar os decisores na formulação de melhores programas.

A construção da teoria da mudança, na linha da avaliação realista, deve ser participativa e suscitar os entendimentos e as teorias implícitas dos stakeholders. A CE sintetiza cinco aspetos que devem ser definidos em conjunto entre o avaliador e os stakeholders para a construção da teoria da mudança (CE, 2013):

- 1 Visão de longo prazo de um programa. é provável que se relacione com uma escala de tempo que se situe para além do seu horizonte temporal. é provável que se relacione com uma escala de tempo que se situe para além do seu horizonte temporal. O objetivo está relacionado com um problema local, regional ou nacional.

² A teoria do programa pode ser utilizada de modo isolado, incluindo os *inputs*, o quadro organizacional, as atividades do programa (modo como o programa é executado), e os resultados intermédios (cadeia de respostas que as atividades desencadeiam e que conduzem a resultados finais).

- Uma vez acordado o objetivo final do programa, as partes interessadas são encorajadas a
- 2 considerar os resultados necessários que serão exigidos até ao final do programa, para que tal objetivo possa ser atingido a longo prazo.

-
- Passos 3 e 4: Pede-se aos interessados que articulem os tipos de resultados a curto prazo que os
- 3 ajudarão a atingir os objetivos especificados. Nesta fase, os envolvidos no programa consideram as atividades ou intervenções mais apropriadas necessárias para provocar a mudança necessária.

-
- Os intervenientes consideram os recursos que podem ser realisticamente utilizados para as
- 5 intervenções planeadas. Estes incluem pessoal e capacidade organizacional, a existência de redes e instalações de apoio, bem como capacidade financeira.

A clarificação de conceitos, segundo a CE, deve ser feita em confronto com a realidade prática, em conjunto com os *stakeholders*, e no momento da aplicação da metodologia, identificando-se os aspetos críticos da execução do programa em prol da avaliação mais adequada (CE, 2013).

A teoria da mudança (Astbury and Leeuw, 2010) – pode ser desenvolvida a partir de diferentes métodos: análise de dados (análise de trajetórias / *path analysis*); modelagem causal / *causal modeling*; observação do programa em ação; entrevistas com funcionários e *stakeholders*; exercícios de mapeamento de conceitos; análise formal de documentos de programas e políticas; pesquisa do estado da arte – com o objetivo comum de criar uma narrativa explicativa de como o programa funciona, com quem e em que circunstâncias.

Existem vários tipos de resultados a considerar (Coryn et al., 2011): realizações (*outputs*) que decorrem diretamente de uma ação, por exemplo, nº de participantes numa formação, nº de empresas apoiadas; resultados (*outcomes*), ou mudanças previstas que ocorrem direta ou indiretamente dos inputs, das atividades e das realizações da intervenção. Os resultados incluem resultados iniciais (e.g. mudanças nos conhecimentos, aptidões, capacidades e outras características), resultados intermédios (e.g. mudanças de práticas) e resultados a longo prazo (correspondem em geral à finalidade das políticas).

Os avaliadores podem formular uma teoria plausível do programa a par de diferentes fontes (Coryn et al., 2011): teorias e investigações pré-existentes (e.g., nas ciências sociais), teorias implícitas (e.g., visões dos *stakeholders*), observações do programa em funcionamento (teoria emergente), e, nos melhores casos, uma combinação das anteriores (teoria mista/integrada).

Chen (2006), autor precursor da TdM, embora com presença discreta nos guias da CE, contribui para este debate, sublinhando a necessidade de o avaliador estabelecer as suas próprias normas e critérios para escolher o método; bem como de desenvolver sistematicamente estratégias de utilização de métodos mistos. A decisão pelo método é uma decisão técnica, que compete à equipa de investigação.

Caixa negra

A CE defende que a teoria da mudança é uma abordagem teórica de programa, que se preocupa em abrir a caixa negra e ir além das descrições dos inputs, procurando compreender as teorias

dos atores no que diz respeito às intervenções dos programas, e por que razão devem funcionar. A clarificação destas teorias é uma das funções do avaliador *ex ante*. Esta é uma figura de estilo frequentemente utilizada em teoria da avaliação que visa descrever o espaço entre o input e o output esperado de um programa (Stame, 2004). As avaliações baseadas na teoria ajudaram a abrir a caixa negra dos programas. O núcleo duro da discussão em torno da TdM é a sua relação com a “caixa negra”, mas diferentes abordagens teóricas tratam o problema da caixa negra a partir de ângulos diferentes.

Para Weiss (1995), a caixa negra contém muitas teorias. A autora chama-lhes "teorias de mudança". Assumem a forma de pressupostos, entendimentos tácitos – mais do que um para o mesmo programa. Todas estas "teorias" têm de ser trazidas à luz a fim de se chegar a um consenso sobre as quais merecem ser testadas (Weiss, 2000). As teorias de mudança têm duas componentes: a teoria da implementação, que define, de forma descritiva os passos a dar na implementação do programa. Centra-se na forma como a intervenção é organizada para chegar aos resultados desejados; e a teoria programática, o núcleo da avaliação, baseada nos mecanismos que fazem as coisas acontecerem. Os mecanismos intervêm entre a prestação do serviço de programa e a ocorrência de resultados. O mecanismo da mudança não é as atividades do programa em si, mas a resposta que as atividades geram (Weiss, 1997). O ABT deve tornar estes mecanismos claros, decompor um programa nos seus mecanismos subsequentes, e utilizar dados de diferentes tipos para os testar. Desta forma, a ABT ajuda a gerar novas teorias.

Não obstante, se para Weiss a caixa se encontra cheia, para outros autores, como Chen e Rossi (1989) a caixa negra encontra-se vazia e não contém nenhuma teoria, já que as ideias tácitas não são claras. A solução, segundo os autores, reside no conhecimento das ciências sociais: as avaliações orientadas pela teoria deveriam fornecer a tais programas uma boa teoria das ciências sociais.

Análise generativa

Segundo Pawson & Tilley (1997), o problema da caixa negra reside no facto de não podermos dizer se a mudança pode ser atribuída ao programa. A abordagem realista proposta pelos autores baseia-se numa teoria ‘generativa’ da causalidade: não são os programas que fazem as coisas mudarem, são as pessoas, inseridas no seu contexto, que, quando expostas aos programas, fazem algo para ativar determinados mecanismos, e mudar. Assim, o mistério da caixa negra é desvendado: as pessoas habitam-na. Isto faz com que o desenho da avaliação seja completamente diferente. O avaliador elabora o modo como o mecanismo poderia funcionar num dado contexto e pede às pessoas que forneçam provas.

A lógica de inferência causal generativa assenta na abordagem metodológica da AbT (contextos + mecanismos = resultados) e articula-se com vários métodos, como a análise lógica, a análise de contribuição, e o *process tracing*; a avaliação realista e a análise de congruência. Pode ainda apoiar-se em variadíssimas técnicas, como a análise documental, análise estatística, entrevistas, inquéritos, observação participante, análise de conteúdo, workshops, *focus groups*, estudos de caso, simulação computacional) (EC, 2008; 2013; Stern et al, 2012; Gates & Dyson, 2017).

Análise de contribuição

A CE cita Mayne (2001) para acrescentar um aspeto relevante relativo à causalidade (Befani & Mayne, 2014); uma teoria da mudança é um modelo de intervenção como causa contributiva; é um modelo do pacote causal que mostra a contribuição da intervenção para os resultados (Mayne, 2012).

As atividades da intervenção raramente são a única causa de um resultado. A teoria da mudança deverá, por isso, descrever um pacote de atividades que se ligam em cadeias causais e os seus pressupostos que, em conjunto, são suficientes para contribuir para os resultados pretendidos.

Cartwright e Hardie (2012) chamam a estes pressupostos fatores de apoio, os eventos e condições necessárias para contribuir para o efeito de uma causa. As atividades de intervenção são essenciais e uma parte necessária deste conjunto causal suficiente. Sem as atividades de intervenção a realização dos pressupostos do nexo de causalidade não seria suficiente para dar uma contribuição. As atividades de intervenção são uma causa contributiva para os resultados.

(Mayne 2012) discute as causas contributivas e os pacotes causais no contexto das teorias da causalidade, e em particular as causas INUS: Insufficient, mas parte necessária de uma condição que é em si mesma desnecessária, mas Sufficient para a ocorrência do efeito (Mackie, 1974).

A teoria da mudança é um modelo da contribuição e não causa per se do resultado pretendido, porque pode haver outros fatores externos que também contribuem para os resultados pretendidos. As influências externas podem ter efeitos positivos ou negativos sobre o nível de resultados alcançados. As influências externas podem explicar parte ou todos os resultados observados (Mayne, 2015).

FEEI

A governação multinível convoca uma caixa negra colossal, com muitas formas diferentes de a preencher, suscitando dois problemas, típicos de relações micro, meso e macro, para os quais têm de ser trabalhadas teorias explicativas (Stame, 2004):

- a) Como atribuir as mudanças ocorridas a nível global (na pontuação dos indicadores de objetivos) à eficácia do programa?
- b) Como atribuir os resultados a determinados parceiros, em detrimento de tantos outros atores envolvidos numa só ação?

Para avaliar a complexidade da governação multinível, Stame (2004) sugere que se recorra à avaliação orientada para a teoria. Contudo, o potencial da ABT não parece ser sido suficientemente explorado no âmbito dos programas apoiados pelo Fundo Europeu de Desenvolvimento Regional (FEDER)³, nos Estados membros, pelo menos até 2013, porque os

³ O FSE parece não apresentar as mesmas dificuldades metodológicas.

programas raramente articulam uma teoria clara da mudança e as avaliações não se baseiam em quadros teóricos rigorosos integrando e orientando a utilização de vários de métodos (Riché 2013).

O objetivo da CE a longo prazo é desenvolver atividades de avaliação da política regional que capitalizem conclusões retiradas de avaliações conduzidas em diferentes contextos e construam conhecimento sobre mecanismos subjacentes semelhantes que conduzem a regularidades de resultados. Essa base de conhecimento descreveria também as alternativas existentes que funcionam em diferentes contextos, favorecendo a adaptação de políticas quando a replicação de políticas não é possível (Riché 2013). O resultado deste processo seria a identificação de "teorias de médio alcance", uma abordagem à teorização sociológica desenvolvida por Robert K. Merton (1968), destinada a integrar a teoria e a investigação empírica. O termo teoria de médio alcance não se refere a uma teoria específica, mas a uma abordagem à construção de teorias.

Ao projetá-las, as Autoridades de Gestão devem explicar ex ante a sua teoria da mudança: como a mudança esperada irá contribuir tanto para os objetivos locais como europeus (teoria do programa) e como as atividades planeadas conduzirão aos resultados esperados (teoria de implementação) (Riché, 2013).

2. A TdM na Perspetiva da Comunidade de Práticas

O presente exercício assenta no pressuposto de que as dimensões teóricas e metodológicas da TdM são indissociáveis da realidade empírica em que se inserem, neste caso, os modos de governação dos Fundos Europeus Estruturais de Investimento – FEEI em Portugal. A avaliação torna-se, assim, objeto de análise dos atores de avaliação que integram a comunidade (ou ecossistema – Barroso, 2022) dos FEEI.

A partir de entrevistas a avaliadores e contratantes, procurou-se recolher perceções sobre as condições de aplicação da TdM na avaliação do PT2020, na expectativa de contribuir para a compreensão dos limites e potencialidades deste modelo metodológico neste contexto específico.

As entrevistas realizadas permitiram-nos transpor o “problema da avaliação” para o “problema da comunidade de avaliadores/as”, nos mesmos termos do exercício já iniciado desde os anos 1960, de transferência do enfoque de análise dos estudos sobre ciência, do “conhecimento” para a “comunidade científica”. Procuramos assim fazer uma socioepistemologia (Jesuino 1993) para compreender o campo de conhecimento a partir da comunidade que o produz (Jesuino, 1995).

A análise qualitativa baseou-se em entrevistas semiestruturadas (N15). As entrevistas incidiram sobre perceções em torno da aplicação da TdM à avaliação dos fundos europeus, veiculadas por atores de avaliação – contratantes e contratados – do “triple hélix” (Etzkowitz & Leydesdorff 1998): administração pública, universidades e empresas – do modelo de governação do Portugal 2020.

Em termos gerais, seguimos Etzkowitz e Leydesdorff (1998), que propuseram o modelo de uma Tripla Hélice de relações universidade-indústria-governo para explicar os desenvolvimentos estruturais nas economias baseadas no conhecimento. O nosso campo empírico corresponde à Tripla Hélice composta pela administração pública (AD&C, Autoridades de Gestão, PlanAPP), empresas de consultoria, e universidades.

As perceções dos entrevistados foram analisadas do ponto de vista da homogeneidade da população, assente no facto de integrarem a comunidade de avaliação dos fundos europeus. A reduzida extensão da população entrevistada e a elevada heterogeneidade dos perfis individuais que a compõem⁴ não nos permitiria explicar a diferenciação das perceções de modo estratificado – sabemos, no entanto, que a explicação por Triple Hélix não singraria, já que a administração pública, aqui representada pela AD&C, a PLANAPP e as Autoridades de Gestão, não são uníssonas no seu posicionamento face à TdM, tendo as Autoridades de Gestão visões por vezes mais próximas das empresas de consultoria – portadoras das críticas mais aguerridas – do que dos restantes representantes da AP.

⁴ Que para além das variáveis supramencionadas beneficiam de uma antiguidade na profissão, posição hierárquica, entidade empregadora, setor público/privado de inserção, díspares, etc.

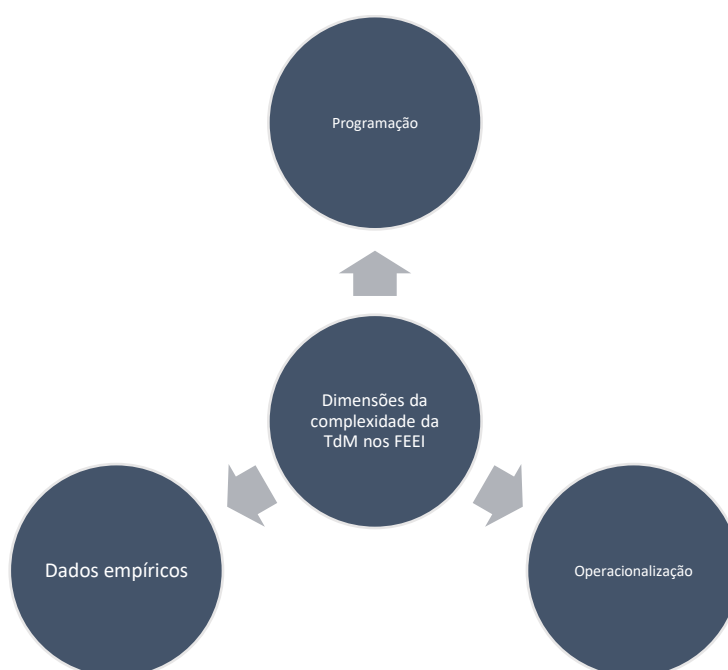
Não foram diferenciados níveis hierárquicos (técnico, chefia, líder) para salvaguarda do anonimato dos entrevistados, que seriam facilmente identificados numa comunidade de avaliação dos fundos europeus de reduzida dimensão.

As entrevistas foram realizadas em 2022, após a transição legal do PT2020 para o PT2030. As gravações foram efetuadas via Microsoft Teams/ZOOM (duração média de 1 hora), transcritas pelo software TRINT, e analisadas via MAXQDA, ao abrigo do Regulamento Geral de Proteção de Dados (GDPR) – Reg. 2016/679.

Discussão dos resultados

De acordo com a pesquisa empírica realizada, a aplicação do modelo da TdM no âmbito da estratégia de avaliação conheceu no PT2020⁵ uma significativa complexidade. A explicação articula-se em 3 dimensões principais: a programação dos FEEI; o modo de operacionalização da abordagem metodológica da TdM; e as dificuldades sentidas, os limites da dimensão empírica desta abordagem (figura 1).

Figura 1. Dimensões da complexidade empírica da TdM aplicada ao PT2020



⁵ Decreto-Lei n.º 137/2014 de 12 de setembro.

Programação

Os programas complexos visam criar sinergias que produzem resultados que iniciativas isoladas não produzem individualmente (Stame, 2004). A programação dos FEEI, altamente complexa, ramifica-se em múltiplas dimensões. Os atores de avaliação têm dificuldade em articular e selecionar esta multidimensionalidade para a construção da TdM: acordo de parceria, domínio temático, objetivo temático, fundo estrutural, prioridade de investimento, indicador de realização (ações), objetivo de intervenções, tipologia de intervenção, tipologia de operação, programa, eixo prioritário, objetivo específico, indicador de resultado, concurso.

De acordo com as entrevistas realizadas, a complexidade dos programas tem consequências diretas na dificuldade em construir teorias da mudança: “A complexidade dos programas é muito grande (...) Estamos a fazer teorias da mudança de programas cada vez mais complexos. É infernal”. (Avaliador 2, Empresa).

Operacionalização

A operacionalização da TdM levanta várias questões aos nossos entrevistados, (i) ao nível da articulação da TdM com a programação, (ii) do momento certo para a sua construção, (iii) da identificação de indicadores para a sua operacionalização, (iv) e em última instância, de crença na utilidade da abordagem metodológica da TdM pelos atores de avaliação e stakeholders:

- i. A coerência entre a teoria do programa e a TdM nem sempre é assegurada: “A teoria de mudança não casa muito com as questões de avaliação. (...) Isso gera uma complexidade tal que temos os relatórios iniciais a serem pedidos com sucessivas revisões. Gera também uma conflituosidade grande entre as equipas de avaliação e a unidade de avaliação da Agência” (Avaliador 1, Empresa).
- ii. Apesar de a CE propor a construção de TdMs a montante do processo de programação, numa fase ex ante, tal não foi possível no PT2020, o que terá gerado uma dificuldade acrescida à sua implementação: “Talvez não estejamos a utilizar a mais-valia da TdM porque ela não está a ser usada na fase de planeamento, não é? (...) Depois na fase de avaliação, acaba por se confrontar com algumas dificuldades ou limitações porque estamos a avaliar um programa que, se calhar não tem um racional tão claro como deveria ter” (Avaliador, Administração Pública).
- iii. Tal como referido na revisão da literatura: “há uma grande dificuldade em definir resultados e a sua quantificação. Qual é o indicador? No Fundo Social é mais fácil, mas no FEDER muito difícil. Acaba por ser um exercício um bocadinho fictício. Nós fixamos indicadores, contratualizamos com os beneficiários. E depois? Portanto, eu não acho que seja um instrumento de gestão real. É um proforma”. (Contratante, Administração Pública).
- iv. Mas o maior entrave à operacionalização da TdM prender-se-á com a desvalorização desta abordagem: “A ADC é muito exigente, mas depois as autoridades de gestão nem todas dão a importância...”. (Avaliador 2, Empresa). Na mesma esteira: “As autoridades de gestão, quando são programas regionais, ficam completamente sem perceber o que

aquelas pessoas estão a discutir, porque não percebem o que está a ser discutido”. (Avaliador 1, Empresa) Coloca-se, deste modo, uma questão de utilidade: “Mas o que é preocupante nisto é que a utilidade para quem recebe a avaliação é cada vez menor, porque passa imenso tempo a discutir coisas que não são os pontos essenciais que eles querem ver. A AD&C passa imenso tempo a discutir metodologias. E os avaliadores passam imenso tempo nisto” (Avaliador 3, Empresa). A questão da utilidade parece implicar os vários stakeholders: “Às vezes há reuniões em que estão os programas todos. (...) todos os PO regionais do Continente, o Compete, o IAPMEI, o Turismo, a ANI e a AICEP, representadas ao mais alto nível. Na reunião houve uma parte que foi só discussão da teoria da mudança. Se focava no impacto. Se aquilo era um pressuposto ou um risco. Aquelas pessoas não querem saber disso, aquelas pessoas querem respostas para os seus principais problemas e não estão para discutir aquilo” (Avaliador 4, Empresa). São ainda colocados em perspetiva objetivos e competências distintas no processo de avaliação: “Mas não temos ninguém que tenha uma competência específica naquela situação (...) depois temos um monólogo, a ADC a discutir questões conceituais, metodológicas com uma equipa técnica. E temos eventualmente aquilo que é o cliente, que é o tomador da avaliação, está-se nas tintas e a borrar para aquilo. Precisa é de um conjunto de recomendações e conclusões que cheguem em tempo útil. Não está para suportar aquela discussão, nem sequer a compreender” (Avaliador 1, Empresa).

Dados empíricos

A avaliação de impacto baseada na teoria produz uma estimativa qualitativa dos impactos (CE, 2018). No entanto, a exigência acrescida imposta à dimensão empírica da TdM (questionários e entrevistas) tem-se tornado por vezes contraproducente no que toca ao rácio fasquia/qualidade: “a família de avaliações baseadas na teoria é de uma complexidade muito grande que vem de mercados muito fortes de avaliação, o americano, o inglês etc. Quando transposta para o mercado nacional, depois é difícil, porque de facto não nós não temos evidência suficiente para suportar muita daquela coisa. É um é um grande desafio metodológico às equipas”. (Avaliador 2, Empresa). Dito de outro modo: “A questão da teoria da mudança e da questão da recolha das evidências que suportam a teoria da mudança não tem sido feita de uma forma proporcional àquilo que é a capacidade de recolha de informação. (...) Ora, qual é o problema dos processos de auscultação? (...) Nós temos duas tendências um bocadinho contraditórias que conflituam entre si, uma pressão para uma recolha cada vez maior de aspetos de recolha de evidências que suportam a verificação de um conjunto de riscos e pressupostos e, ao mesmo tempo, uma disponibilidade cada vez menor dos beneficiários responderem a questionários. (...) Simplesmente nós não podemos querer ter tudo”. (Avaliador 1, Empresa). Esta questão desemboca numa outra: “E há outra falácia neste processo que é nós estarmos a fazer a teoria da mudança passado três ou quatro anos da sua programação, em que, com a dinâmica que as organizações públicas têm, já não é possível juntar à mesa as pessoas que estavam naquele momento, nós juntamos as entidades, mas não as pessoas” (Avaliador 4, Empresa).

Os limites empíricos não se circunscrevem à nossa comunidade de avaliação. A questão da qualidade dos dados empíricos é uma questão antiga nos FEEI. De acordo com Hart (2007) a comunidade de avaliação dos estados-membros tem dificuldade em aceder a dados precisos (geralmente quantitativos) sobre os quais aplicar as metodologias existentes. O autor sugere a adoção de metodologias mais apropriadas, igualmente exigentes em termos de dados, mas de forma diferente.

A este propósito, a CE determina para o PT2030 (DG Regio 2021):

- Melhores avaliações: Evitar fazer demasiadas perguntas de avaliação para garantir que a avaliação se centre nas questões políticas críticas, sobre as quais são necessárias provas, identificadas em diálogo com as partes interessadas e com os decisores político. Os objetivos de monitorização e os indicadores de desempenho devem ser da responsabilidade dos sistemas de monitorização.
- Alargar a utilização de dados: Utilizar dados administrativos sempre que possível para avaliar os efeitos das intervenções. Deve ser dada especial atenção ao tempo de armazenamento dos registos de dados públicos que abrangem amostras ou populações totais de entidades. Se forem necessários dados específicos, a gestão pode ter de negociar com a autoridade competente (por exemplo, serviço nacional de estatística) com antecedência para assegurar que os dados estão disponíveis.
- Selecionar os métodos de avaliação corretos: Evitar confiar apenas em informações provenientes de inquéritos e/ou entrevistas para identificar os efeitos de uma intervenção e verificar a sua validade através de triangulação. Os estudos de casos podem ser importantes e dar uma visão da influência dos mecanismos através dos quais as medidas a serem examinadas alcançam os seus efeitos. Contudo, os estudos de casos precisam de ser aprofundados, a uma escala representativa. Ao adotar uma abordagem baseada na teoria, explicita em pormenor a cadeia causal através da qual se espera que uma intervenção conduza aos resultados desejados, investigue se os vários elos causais são evidentes na prática e modificam a 'teoria' quando necessário à luz de descobertas empíricas.
- As avaliações de programa que analisam a "teoria da mudança" ocorrem sobretudo numa fase posterior, quando houver provas (execução) suficientes para se tirarem conclusões.

Corolário

Apesar das críticas feitas pelos atores de avaliação ao modo como a TdM foi implementada no seio da estratégia de avaliação do P2020 e no âmbito do triple hélix (Administração pública, universidades, empresas) todos os entrevistados consideram que a TdM é um passo irreversível, ou mesmo um salto epistemológico: “Eu acho que isso foi um passo crucial. Acho que é um passo decisivo e eu atrevo-me a dizer que é um salto epistemológico, desde logo porque acho que é muito mais fácil identificar vendedores da banha da cobra quando se obriga as pessoas a pensar as coisas nestes termos”. (Avaliador, Universidade). Num registo similar: “Há um salto

qualitativo da parte da agência, naquilo que entende que se deve ter, as teorias de mudança e na construção a partir das teorias de mudança”. (Avaliador 1, Empresa). No mesmo sentido, e de modo fundamentado: “Acho que há duas formas que podemos ver a teoria a mudança. Por um lado, o que é a programação, ou seja, do lado de programação dos vários períodos de programação e mais uma vez aí temos sempre evoluções qualitativas do QCA para o QREN e do QREN para o PT2020. Em termos de programação, também há aí uma evolução muito grande do que é solicitado em termos de programação. Nós no Portugal 2020 já temos muitos elementos. Não diria, não diria da teoria da mudança, mas diria do quadro lógico, na própria programação, que a programação já incorpora um conjunto de elementos do quadro lógico, muito bem. Não quer dizer que não existissem antes, mas eram mais difusos. No PT2020 eles estão muito, muito presentes. Sabe-se perfeitamente qual é o problema, quais são os meios, quais são as realizações, os resultados e os impactos. Está muito bem definido. E quais são os principais indicadores que vão medir esses elementos⁶” (Contratante, Administração Pública).

⁶ Se esta posição contradiz em parte opiniões de entrevistados citadas anteriormente, tal como enunciado no ponto metodológico, a análise foca-se na validade cumulativa das perceções, sem se centrar na estratificação da população e na diferenciação de posicionamentos face à TdM.

3. A caracterização das intervenções complexas e os desafios que coloca à avaliação

O principal contributo das estratégias de avaliação baseadas na teoria é o de ajudar a responder às seguintes questões: “o que funciona melhor, para quem, em que circunstâncias e porquê (Pawson and Tilley, 1997)?” Deste modo, espera-se da avaliação um contributo acrescido para o objetivo de melhorar e tornar mais eficazes as políticas públicas. A dificuldade de resposta àquelas questões é tanto maior quanto maior for a incerteza que caracteriza as intervenções e a sua permeabilidade ao contexto e dinâmica de interação em que inscreve a sua ação. Ou seja, a complexidade das intervenções acrescenta dificuldade na resposta às questões que motivam a procura de novas abordagens metodológicas e, em particular, na proposta de abordagem metodológica baseada na Teoria da Mudança (TDM).

Como referem Connel e Kubisch (1998), embora os passos para articular uma TdM sejam similares, essa semelhança desvanece-se quando as intervenções têm múltiplos domínios de intervenção e operam em diferentes níveis. É esse o caso das intervenções abrangentes de âmbito comunitário que estudaram no seu trabalho. Como bem explicitam, essas intervenções além de situarem múltiplos domínios de intervenção, envolvem um grande conjunto de atores na sua conceção e implementação e evoluem (transformam-se) ao longo do tempo.

Os mesmos autores referem, ainda, que a construção de uma boa TdM é essencial à avaliação deste tipo de intervenções. Esta deve ser: plausível, fazível (na ótica da programação) e testável (na ótica da avaliação). Para a construção de teorias da mudança com estas propriedades é fundamental combinar diversas fontes de informação e saber articulá-las de forma a conseguir uma imagem fidedigna do processo de mudança procurado pelas intervenções de âmbito comunitário.

A leitura que releva o papel da natureza e perfil das intervenções para definir as estratégias de avaliação é sublinhada por muitos outros autores. O perfil das intervenções pode ser analisado de acordo com a seguinte tipologia (Rogers, *et al*, 2015): simples, complicadas ou complexas. Os aspetos simples das intervenções correspondem a objetivos específicos e bastante padronizados. Os aspetos complicados referem-se a domínios de atuação com múltiplas componentes que integram um amplo quadro de relações causais. Os aspetos complexos referem-se a dinâmicas emergentes das intervenções, ou seja, a mecanismos de adaptação e reação da necessidades e oportunidades emergentes. Retomaremos adiante de modo mais sistemático esta distinção.

Esta proposta de tipologia releva para sublinhar que o planeamento da avaliação depende da natureza do objeto a avaliar e que este apresenta complexidades diferenciadas. A mobilização da TdM é particularmente sensível a este tipo de propriedades das intervenções, na medida em que procurar estabelecer e compreender o modo de funcionamento das relações causais inerentes a uma determinada intervenção, exige abordagens diferentes em função da natureza dos programas. Ora, responder a isto num contexto de complexidade é mais difícil. A caixa negra inerente a intervenções complexas é maior e, nessa medida, mais exigente o processo de as compreender.

São várias as abordagens que na literatura é feita à complexidade. Numa ótica mais genérica, a complexidade pode ser entendida como refletindo sistemas cujas características tornam difícil compreender e prever o seu comportamento (Barbrook-Johnson, 2020). Ora, tornando-se difícil compreender o funcionamento destes sistemas torna-se, conseqüentemente, igualmente difícil avaliar os resultados da sua ação. São diversos os atributos referenciados a estes sistemas pelo autor: natureza dinâmica e adaptativa, múltiplos níveis de funcionamento, sem delimitação e com fronteiras pouco definidas, combinando domínio de estabilidade com instabilidade. A imprevisibilidade emerge como atributo central destes sistemas e coloca, por consequência, uma acrescida dificuldade metodológica para a sua avaliação. Ora, quando as intervenções se integram em contextos sistémicos complexos ou são elas próprias marcadas por atributos de complexidade, os desafios que se colocam à prática de avaliação são maiores.

A compreensão do conceito de complexidade e dos atributos que a expressam surge como condição fundamental para a sua apropriação pela avaliação. Conforme refletido por Barbrook-Johnson (2020), o entendimento do conceito de complexidade entre os avaliadores é muito diverso e pouco propício à sua mobilização – da perspetiva da complexidade – pelos avaliadores. A complexidade deve ser considerada como uma oportunidade no contexto da avaliação e convocar uma abordagem mais pragmática no que diz respeito às técnicas e métodos de avaliação. Desse modo, será viável situar o que é possível e o que não é no contexto de exercício específicos de avaliação e as suas implicações. Nesta perspetiva, o entendimento da complexidade das intervenções deve estar presente quando a avaliação é encomendada e, desse modo, apoiar a determinação dos seus requisitos e abordagem metodológica.

A utilização da abordagem baseada na Teoria da Mudança constitui um recurso útil à abordagem da complexidade na avaliação de instrumentos de política pública. Com efeito, a Teoria da Mudança permite formalizar a ótica subjacente ao planeamento das intervenções e convocar para a avaliação o racional subjacente à lógica da intervenção. A compreensão desta dinâmica apoia o mapeamento da estratégia de ação e dos resultados que razoavelmente se podem projetar.

Outro importante recurso associado à TdM é o de contribuir para integrar a ótica de avaliação no process político (Barbrook-Johnson, 2021). Isto é, considerar a perspetiva da avaliação no momento da identificação de prioridades políticas, de desenho das políticas e do seu planeamento, contribuindo, também, para promover abordagens avaliativas ex-ante. Deste modo, a condução de exercícios de avaliação é acautelada e melhor preparada, sobretudo quando se trate de intervenções complexas, na medida em que é possível garantir uma maior sistematização dos recursos e condições necessárias à condução da avaliação.

De referir que é crescente, também, a tendência para que a iniciativa política defina intervenções que integram múltiplos projetos, *stakeholders* e modos de governação, exigindo um esforço maior de planeamento e um maior investimento na etapa de desenho do exercício de avaliação.

Destas considerações resulta a perspetiva da importância que tem o processo de desenho e encomenda da avaliação para a capacidade de a integrar e lhe dar resposta na prática. A flexibilidade na definição e condução da abordagem metodológica e a promoção de um diálogo

continuo e aberto são recursos relevados na resposta à avaliação de instrumentos de política mais complexos

Importa, assim, discutir o que constitui um campo complexo de avaliação. A discussão da complexidade das intervenções associa, comumente, uma outra que é da natureza complicada das mesmas. Ou seja, complicação e complexidade surgem como graus de sofisticação das intervenções (sistemas) e ambas colocam uma maior exigência ao exercício da avaliação.

Rogers (2008) convoca a literatura precedente para situar a distinção entre intervenções complicadas e complexas. De forma mais simplificada, as primeiras situam múltiplas componentes, tanto ao nível da operacionalização e da governação como da causalidade implícita, fazendo emergir a necessidade de considerar essa diversidade no contexto da avaliação. No caso dos segundos, as intervenções complexas, consideram-se, adicionalmente, a evolução continua das intervenções através de processos internos de mudança e dinâmicas emergentes (não previstas) em resultado da própria ação (causalidade recursiva). Neste caso, as intervenções transformam-se ao longo do tempo.

Por oposição a estas realidades, poderemos considerar que intervenções e sistemas simples são aqueles que se caracterizam por uma maior circunscrição do âmbito de intervenção, maior linearidade na cadeia de causalidade, uma natureza mais prescritiva da ação e um quadro de participação mais delimitado. Nestes casos, o contributo da abordagem baseada na teoria para (re)construir a lógica da intervenção e o seu quadro de funcionamento não perde utilidade, mas será menos necessária.

É a intervenções deste tipo que melhor se aplicam os modelos lógicos de programação que situam de forma linear a causalidade entre recursos, atividade, realizações, resultados e impactos. Porém, como já vimos, a produção dos resultados esperados e dos consequentes impactos pode, em contextos de ação mais elaborados (complicados e complexos) ser influenciada por dimensões relacionadas com o contexto de implementação, a interação com outras intervenções, as características dos públicos e a própria retroalimentação da programação gerada no contexto da intervenção. Nestes casos, os quadros lógicos de programação caracterizam-se por uma menor linearidade da articulação dos seus elementos de programação e na previsibilidade dos seus resultados.

Ainda recorrendo a Rogers (2008) é possível sinalizar três atributos essenciais que distinguem as intervenções complicadas. São eles:

1. A existência de múltiplas entidades de governo das intervenções, quase sempre com diferentes âmbitos de atuação e de jurisdição. Este atributo contrapõe-se aos modelos de governação simples protagonizados por uma única organização.
2. A existências de múltiplas cadeias de programação, ou seja, a intervenção convoca múltiplas cadeias de causalidade na produção dos resultados visados. As intervenções simples situam, comumente, uma cadeia de causalidade.
3. A existência de mecanismos de causalidade alternativos, isto é, que podem operar em diferentes contextos e circunstâncias. Este atributo confere às intervenções a capacidade de se moldarem a contextos de intervenção diferenciados. No caso das

intervenções simples a intervenção programada é universal, aplicando-se ao contexto de ação visado pela intervenção.

No caso das intervenções complexas, os principais atributos identificados são:

1. A existência de uma causalidade recursiva com várias interações circulares, ou seja, as cadeias de causalidade produzem novas e diferentes dinâmicas de causalidade que criam condições de sucesso. Adicionalmente, os efeitos gerados podem ser significativamente ampliados com limitados esforços incrementais. Esta dinâmica contrasta com a estabilidade e linearidade das cadeias de causalidade simples em que um determinado estímulo produz um certo resultado.
2. A emergência de novas dinâmicas de interação que produzem mecanismos e resultados não previstos. Esta dinâmica reflete a adaptação das intervenções a novos contextos.

Esta sistematização permite reconhecer que uma significativa parte das intervenções no âmbito da política pública reúnem atributos de complexidade (aqui usada como conceito agregador) que conferem exigência acrescida ao desenho metodológico da avaliação e à sua implementação.

Recorrendo à classificação proposta por Rogers (2008)⁷ para distinguir a natureza das intervenções, analisa-se em seguida o nível de complicação e complexidade atribuível aos instrumentos programáticos que costumam mediar a aplicação dos fundos europeus e ser objeto dos estudos de avaliação comissionados. Mobilizamos para o efeito a três tipologias de intervenções comumente sujeitas a exercícios de avaliação no âmbito dos fundos europeus, estruturando a análise em torno dos atributos acima referenciados.

Tabela 1. Os Atributos das Intervenções Complexas e os Instrumentos de Política dos FEEI

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
Governança e Implementação	Os modelos de governação implementados mobilizam vários atores numa estrutura de gestão com vários níveis e responsabilidades.	Os modelos de governação implementados mobilizam vários atores numa estrutura de gestão com vários níveis e responsabilidades. A dimensão regional da	Os modelos de governação implementados mobilizam vários atores numa estrutura de gestão com vários níveis e responsabilidades.

⁷ Rogers, Patricia J. (2008), "Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions", SAGE Publications, London, New Delhi and Singapore.

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
		intervenção alonga a cadeia de governo.	Por darem expressão à governação multinível a complexidade da estrutura de governação e implementação é acrescida, combinando três níveis de responsabilidade, vários atores e jurisdições territoriais com níveis de sobreposição.
Cadeias Causais Simultâneas	O âmbito das intervenções é muito diverso. Além de agendas setoriais temáticas com alguma especificidade, os programas temáticos acolhem um relevante portfólio de instrumentos de política sendo incontornável a existência de uma multiplicidade de cadeias causais.	O âmbito das intervenções é muito diverso. Além de agendas setoriais temáticas com alguma especificidade, os programas temáticos acolhem um relevante portfólio de instrumentos de política, sendo incontornável a existência de uma multiplicidade de cadeias causais. A natureza regional da abordagem amplia ainda mais o âmbito da intervenção e a diversidade de setores e a abrangência dos impactos visados (vinculada a objetivos de	Embora a seletividade dos instrumentos de política considerados confira uma maior delimitação do âmbito das intervenções, estas acolhem-se em referenciais estratégicos amplos e multiplicam, também, as cadeias causais.

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
		desenvolvimento territorial).	
Cadeias Causais Alternativas	<p>A existência de cadeias causais alternativas pressupõe flexibilidade adaptativa da programação. Definir em que moldes pode ser entendida esta flexibilidade não se afigura linear e beneficia de uma análise da estratégia de implementação de cada instrumento. Contudo, a rigidez que caracteriza a programação dos instrumentos incluídos nos programas operacionais que aplicam os fundos europeus sugere a inexistência deste atributo. A programação regional poderia ser um contexto mais propício à afirmação desta maleabilidade. Contudo, a modelação dos instrumentos a contextos regionais diversos é um progresso inconsistente, assim como a delegação da autonomia necessária nos agentes responsáveis pela implementação.</p> <p>Com efeito, a existência de cadeias causais alternativas pressupõe a possibilidade de diversificar a lógica causal de uma dada intervenção, sobretudo por via da diferenciação das atividades e mecanismos. Constituem cadeias causais paralelas que não estão pré-terminadas, refletindo capacidade de adaptação ao contexto. Esta flexibilidade n</p>		
Causalidade não linear e resultados não proporcionais	<p>A existência de dinâmicas de causalidade que se vão retroalimentando, ou seja, em que a interação entre a implementação e a programação está presente é comum a vários dos instrumentos tipificados. Nesta medida, os resultados alcançados influenciam o contexto e o desenho das intervenções ao longo do tempo, uma transformação incremental e progressiva das cadeias de causalidade. Essa realidade é particularmente expressiva no modo como os instrumentos de política vão sendo afinados em continuo e, com particular significado, como se transformam ciclicamente ao longo dos períodos de programação que organizam a aplicação dos fundos europeus.</p> <p>Menos clara é a associação às mesmas da produção de resultados não proporcionais. A evidência disponível figura-se menos adequada a reconhecer esta dinâmica. Em todo o caso, este atributo de complexidade estará genericamente presente nos programas tipificados.</p>		
Resultados emergentes	A existência de resultados emergentes está presente em muitas das medidas que integram os PO Temáticos. Isso é	Nos programas de base regional e comunitária, a interação da programação com os contextos de intervenção tende a ser mais significativa, contribuindo para uma maior diversificação ao nível dos resultados. Por outro lado, a programação potencia a formulação de projetos	

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
	particularmente evidente em domínios sociais de atuação e naqueles em que a programação é mais aberta e contempla, por exemplo, dinâmicas de capacitação institucional e comunitária. Acresce que em vários domínios a programação é relativamente flexível e imprecisa quanto à dimensão de resultados, encorajando mesmo o contributo de dinâmicas emergentes.	territorialmente contextualizados e com reforçada integração com dinâmicas económicas e sociais territoriais. Estas condições potenciam a formação de resultados emergentes e acentuam a relevância deste atributo no perfil de complexidade destes instrumentos.	

Na tabela seguinte sistematiza-se a análise feita, classificando as diferentes tipologias de programas face aos atributos de complexidade das lógicas de programação e de implementação.

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
Governança e Implementação	Sim	Sim	Sim
Cadeias Causais Simultâneas	Sim	Sim	Sim
Cadeias Causais Alternativas	Sem evidência clara	Sem evidência clara	Sem evidência clara

Atributo	PO Temático	PO Regional	Instrumentos Territoriais
Causalidade não linear e resultados não proporcionais	Sim	Sim	Sim
Resultados emergentes	Sim	Sim	Sim

Como decorre da análise feita, as intervenções operacionais que tipicamente organizam a avaliação no âmbito dos fundos estruturais reúnem elevados níveis de complexidade (são simultaneamente complicadas e complexas). Adicionalmente, o grau de formalização dos quadros lógicos das intervenções e das cadeias de causalidade é genericamente deficitário e limita a explicitação do seu posicionamento face a alguns dos atributos considerados. Estas características interpelam as estratégias de avaliação a adotar na sua avaliação.

A opção pela utilização da teoria da mudança é valorizável pela maior capacidade que confere de explorar analiticamente algumas das dimensões de complexidade referidas. A verificação de cadeias de causalidade simultâneas e alternativas, a verificação dos mecanismos de produção de resultados e impactos e, nessa medida, a compreensão das dinâmicas de interação que reconfiguram os instrumentos e a possibilidade de explorar resultados emergentes são alguns exemplos.

A formalização da teoria da mudança e a sua verificação e atualização em contínuo constitui um recurso para a avaliação. As estratégias baseadas em análises contrafactuais não terão a mesma capacidade de desocultar estas dinâmicas e ajudar a compreender o modo como funcionam os mecanismos e se produzem os resultados e impactos. Do mesmo modo, será mais difícil situar a emergência de novas dinâmicas ao nível da programação. A consistência da abordagem contrafactual é maior quando em causa estão cadeias lineares de programação e efeitos aferidos nos indivíduos. Quando estão em causa a abordagem de nível coletivo (comunitário) o recurso à análise contrafactual tem menor aplicabilidade. Esse é o caso das intervenções consideradas, sobretudo quando pensadas de forma integrada.

Porém, também ao nível da operacionalização da teoria da mudança na avaliação se evidenciam desafios relevantes. Um dos principais desafios tem a ver com a dificuldade em desenhar e aplicar modelos preditivos da mudança a sistemas complexos (Barnes, 2003). Com efeito, a teoria da mudança revela grande aplicabilidade para compreender a lógica da intervenção (o que está subjacente ao seu desenho) e a proposta de intervenção e resultados subjacentes. Nesta ótica, ela permite guiar a verificação das realizações, resultados e impactos de uma determinada intervenção.

Contudo, como refere Barnes (2003) a sua aplicabilidade a sistemas complexos em que o contexto é difícil de definir e captar e em que a interação entre as diferentes dimensões do sistema muda a sua natureza e propósito essa aplicabilidade é mais limitada. Em realidades mais

circunscritas - o nível do projeto, por exemplo - e estáveis, a teoria da mudança constitui um importante referente de aprendizagem e de avaliação. Em realidades mais abrangentes - intervenções de âmbito nacional e comunitário, por exemplo -, a utilização da teoria da mudança enfrente maiores dificuldades. Elas decorrem da evidente dificuldade em predizer e testar lógicas de causalidade a partir de cadeias de impacto formadas em dinâmicas muito interdependentes, mutáveis e agregadas do ponto de vista dos impactos esperados.

O modo de ultrapassar essas dificuldades reúne contributos diversos. A opção por utilizar múltiplas teorias da mudança considerando as dimensões de complexidade é uma delas. Esta abordagem metodológica permite segmentar o exercício de avaliação em segmentos mais homogêneos.

A partir da experiência da avaliação de programas europeus, Stame (2004) releva a utilidade de distinguir a abordagem de nível local e a do programa no seu conjunto. Estes níveis são distintos e sugerem o desenho de propostas teóricas autónomas. No nível mais abrangente o desenho da avaliação pode beneficiar com o desenho de uma proposta teórica que reflita a complexidade da intervenção. Nesta ótica, a abordagem teórica deve ser suficientemente ampla para incluir a diversidade de stakeholders da intervenção, teorias de mudança emergentes no processo e as dinâmicas contínuas de interação. É também este o sentido da reflexão de Barnes (2003). De acordo com esta proposta, a natureza preditiva e estável das teorias da mudança deve ser articulada com o desenho de um modelo lógico mais aberto que permita tomar em consideração a complexidade das intervenções.

A utilização de abordagens metodológicas diversificadas e participativas é uma condição essencial à implementação de modelos lógicos mais abertos e maleáveis.

Contudo, a validação através da evidenciação de consensos entre vários *stakeholders* quanto à teoria da mudança de uma dada intervenção e a subsequente procura de evidências quanto aos efeitos gerados pode afigurar-se insuficiente, designadamente na compreensão de intervenções complexas. Em causa está a capacidade de atribuir causalidade à programação (e ação que dela resulte) subjacente a uma dada intervenção. Dito de outro modo, importa utilizar estratégias metodológicas que permitam compreender como funcionam as intervenções. Para tal é necessário tomar em consideração os atributos e dinâmica de mudança dos contextos em que se desenvolvem.

As abordagens baseadas na teoria procuram assegurar resposta ao desafio da atribuição da causalidade, mas pode haver nelas diferentes expressões metodológicas e, consequentemente, uma eficácia e adequação diferenciada na resposta à análise da causalidade. Valerá a pena situar os termos gerais do debate neste plano uma vez que ele enquadra o desenvolvimento e estratégia metodológica que propomos para este estudo nas etapas subsequentes. O confronto entre os pressupostos metodológicos da Teoria da Mudança e da Avaliação Realista pode ser útil para este propósito.

A abordagem proposta pela Teoria da Mudança procura superar as limitações das abordagens experimentais. A proposta experimental é limitada quando se trata de explicar o como e o porquê e, consequentemente, na resposta mais completa às dimensões avaliação que remetem para a eficácia e eficiência das intervenções. A ótica agregada na análise de resultados a partir

de grupos diferenciados e a ausência de ponderação da heterogeneidade de circunstâncias (contexto) em que se desenvolvem as intervenções emergem como principais limitações.

As abordagens baseadas na teoria enfatizam a relevância do contexto, ou seja, da diversidade de circunstâncias para determinar a causalidade nas intervenções avaliadas. A abordagem realista confere uma especial atenção às condições em que os resultados se produzem e, nessa medida, adequa-se à exigência de olhar que é colocada pela avaliação de intervenções complexas. Ou seja, a abordagem realista confere especial atenção aos mecanismos de produção de efeitos das intervenções. A ênfase colocada pela Avaliação Realista nos mecanismos de produção de efeitos importa para melhor enquadrar o desenvolvimento do presente trabalho.

Recorrendo a Blamey e Mackenzie (2007), podemos situar dois planos complementares de conceptualização da abordagem baseada na teoria:

- O plano que remete para a cadeia de ligação entre objetivos, atividades e resultados visados, considerado como a “teoria da implementação” (Weiss, 1995). Neste plano está em causa definir como os objetivos visados são traduzidos respostas (atividades) e a sua relação com os resultados esperados.
- O plano que remete para a ótica da causalidade, ou seja, o que procurar determinar os mecanismos que sustentam a produção dos resultados esperados. Esta é a ótica explicativa, isto é, a que permite produzir conhecimento sobre como se processam os resultados em determinadas circunstâncias. Weiss (1995) classifica esta dimensão de análise como a teoria da programação.

A importância de contemplar estes dois planos de análise reforça-se no âmbito da avaliação de intervenções complexas. Essa será a forma de resgatar a avaliação de leituras simplificadas da cadeia de relação (objetivos, atividades e resultados) e de prevenir uma compreensão equivocada do funcionamento de intervenções que funcionam em contextos não lineares, caracterizados pela pluralidade de interações e pela mutabilidade das circunstâncias. Trata-se de determinar o como e o porquê em contextos incertos (realidades pouco previsíveis). Nesta medida, afigura-se útil atribuir uma atenção específica a duas abordagens teóricas que procuram apresentar propostas para melhor sistematizar a resposta a estes desafios.

3.1 A Avaliação Realista

A Avaliação Realista confere uma ênfase acrescida a esta proposta analítica. Fá-lo posicionando a análise em torno da configuração de uma cadeia de relação entre contexto, mecanismos e resultado (na sigla inglesa CMA – Context, Mechanisms and Outcome). A abordagem proposta por Pawson and Tilley (1997) sublinha a importância do conceito de mecanismo e convoca-o como forma de dar resposta a necessidade de explicar em que condições (como e porquê) uma intervenção funciona.

A diferenciação da abordagem proposta pela Teoria da Mudança e da Avaliação Realista ajuda a compreender a importância de densificar a abordagem aos mecanismos na avaliação de

intervenções complexas. O quadro seguinte procura sistematizar o modo como conceptualmente se diferenciam esses olhares.

Tabela 2. A TdM e a Avaliação Realista na Avaliação de Intervenções Complexas

Teoria da Mudança	Avaliação Realista
<p>Articulação idealizada da teoria.</p> <p>Estreita interligação entre os planos de implementação e programação.</p> <p>A teoria é validada através da participação de um amplo leque de stakeholders, dando resposta por esta via ao desafio da atribuição de causalidade.</p> <p>Atribuição dos resultados ao quadro lógico desenhado.</p> <p>O desenho do quadro lógico da intervenção apoia-se mais no conhecimento e experiência prévia do avaliador do que na relevância atribuída por aqueles que intervêm no nível de implementação.</p> <p>Apropriação pela comunidade do racional da intervenção.</p> <p>Abordagem cumulativa à construção do conhecimento. Reduzido investimento na interdisciplinaridade na construção do conhecimento (mais confinado aos exercícios de avaliação).</p>	<p>Articulação da teoria com base num processo incremental de produção de conhecimento.</p> <p>A validação é feita através de processos sucessivos de teste e refinamento da proposta teórica.</p> <p>O foco é colocado no teste das cadeias de causalidade que articulam contexto, mecanismos e resultado.</p> <p>Envolvimento menos alargado dos stakeholders</p> <p>Abordagem cumulativa à construção do conhecimento. Acrescido investimento na aprendizagem sobre o modo como os mecanismos funcionam (reforço da ótica interdisciplinar).</p> <p>Depuração e especialização dos processos de aprendizagem e validação.</p> <p>A dimensão explicativa da abordagem tem maior aplicação no plano micro.</p>

A sistematização feita permite sinalizar os desafios que se colocam à abordagem baseada na teoria para a avaliação de intervenções complexas e, no caso em apreço, das que são conduzidas para a avaliação da aplicação de fundos europeus. Com efeito, a complexidade confere utilidade a um maior investimento na compreensão do funcionamento das intervenções e no modo como estas produzem, ou não, os efeitos esperados. Nesta ótica, a ênfase na conceptualização em torno da cadeia contexto, mecanismos e resultados mostra-se mais valorizável.

Esta ótica permite uma melhor ponderação das condições de produção dos resultados e impactos esperados e a leitura das dinâmicas de interação que os determinam (negativa e positivamente). É, também e nessa medida, uma abordagem mais interdisciplinar e propícia à aprendizagem incremental apoiada em sucessivos ciclos de problematização e teste.

Porém, se mais propícia à apreensão da incerteza presente nas intervenções complexas, esta abordagem revela-se menos viável quando em causa estão intervenções de amplo espectro com forte expressão comunitária e escala regional ou nacional. Estas circunstâncias parecem pouco compatíveis com o foco e especialização (progressiva) da análise inerente à abordagem centrada nos mecanismos. Ao invés, a análise baseada na validação de quadros lógicos construídos à partida com base em experiência prévia afigura-se mais maneável. Neste caso, a atribuição de causalidade não poderá ser objeto de uma ênfase tão detalhada na compreensão dos mecanismos.

Duas conclusões lógicas deste breve enunciado conclusivo seriam: a de que a avaliação da complexidade beneficia com uma maior circunscrição do objeto da avaliação; a de que a prática deve privilegiar soluções duais como sugerido por Blamey e Mackenzie (2007) que permitam articular a leitura integrada de intervenções abrangentes com “zooms” analíticos sobre intervenções mais delimitadas.

O roteiro analítico que desenvolvemos neste trabalho permite explorar a reflexão, apoiada em testes práticos, em torno do modo como operacionalizar a avaliação de intervenções complexas. Na síntese conclusiva do trabalho retomaremos a problematização que aqui enunciámos, procurando sinalizar pistas úteis para guiar a prática.

3.2 As TdM “encaixadas”

Considerando as múltiplas dimensões da complexidade reveladas pela análise socioepistemológica da TdM, propomos uma abordagem metodológica que visa racionalizar este modelo.

Esta abordagem inspira-se num tipo de análise de contribuição, as “nested theories” ou “teorias encaixadas” desenhadas por Mayne (2015) e Mayne & Johnson (2015), com o intuito de avaliar políticas complexas com diversas intervenções (figura 1). Neste modelo, a TdM é composta por várias sub TdM, micro e meso, que compõem a TdM macro. A complexidade é compartimentada até perfazer um conjunto global.

Figura 2. TdM – modelo das “teorias encaixadas”

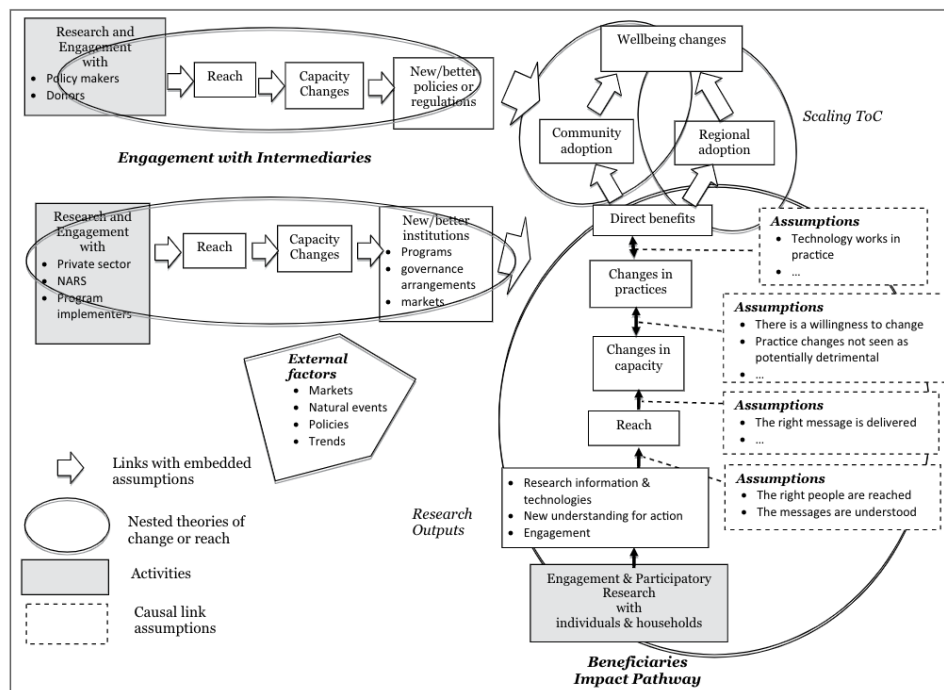


Figure 3. An indicative theory of change for A4NH interventions.

Mayne and Johnson: Theories of change in the CGIAR research program

413

Fonte: Mayne & Johnson (2015)

Para intervenções complexas é assim possível desenvolver uma subteoria da mudança para cada grupo-alvo chave – uma teoria de mudança encaixada ou teoria de alcance. As teorias de alcance podem interagir umas com as outras na obtenção dos resultados desejados. As teorias da mudança mais densas podem ser simplificadas concentrando-se em vias-chave encaixadas de impacto (key nested impact pathways) (Mayne, 2015).

Com base nas “teorias encaixadas” várias teorias da mudança podem perfazer vários retratos sucessivos, autônomos. Estes retratos deverão ser estanques, ter uma autonomia relativa e um caráter complementar e interdependente entre si, perfazendo no seu conjunto uma rede de TdM encaixadas.

A ideia de compartimentação está presente no Plano Global de Avaliação da AD&C quando postula que várias avaliações temáticas perfazem uma avaliação transversal ou global, ou que uma avaliação intercalar é uma meta avaliação. Propomos assim um modo de adaptação da TdM a esta realidade. Uma lógica que permita percorrer os diferentes tipos de avaliação (em função do objetivo, da incidência, dos critérios e do período temporal) e as geometrias variáveis da programação dos fundos.

O desafio de racionalização da complexidade só será eficaz se se proceder à rutura com vários hábitos de avaliação descritos no ponto 2. A TdM não pretende esgotar todas as dimensões da programação, podendo ancorar na seleção de alguns eixos chave. Os atores da avaliação deverão ter a coragem de selecionar/excluir eixos de análise, sejam eles escalas analíticas, escalas temáticas, ou escalas geográficas.















A teoria da mudança não deverá replicar a complexidade da teoria da programação para poder estudá-la, sob pena de não conseguir estudá-la de modo fidedigno. O avaliador deverá ter a ousadia criativa e inovadora de selecionar o princípio que preside à TdM (no âmbito da geometria variável da programação, de modo não exaustivo), sem ter medo de perder legitimidade. Uma teoria da mudança não necessita reproduzir exaustivamente todos os eixos de um programa para o conseguir avaliar. Poderá conter apenas e sobretudo as dimensões axiais relativas ao objetivo da avaliação.

Estão disponíveis vários softwares de representação gráfica da TdM⁸, com diferentes níveis de sofisticação de comunicação visual. Apresentamos uma representação visual em quadro, (permite uma melhor racionalização da apresentação da TdM) que foi efetuada manualmente.

⁸ <https://www.theoryofchange.org/toco-software/>

[https://usaidlearninglab.org/sites/default/files/resource/files/software for visualizing theories of change.pdf](https://usaidlearninglab.org/sites/default/files/resource/files/software%20for%20visualizing%20theories%20of%20change.pdf)

Tabela 3. Teoria da Mudança à escala da TI, abordagem generativa, realista.

<div>  TEORIA DA MUDANÇA: “DESIGNAÇÃO DA POLÍTICA” <i>Como e porquê uma política (não) funciona?</i> </div>	
COMPONENTES	DESCRIÇÃO
PRESSUPOSTOS (Coerência)	 <p>Condições globais para a verificação das ligações causais. Pré-condições, fatores, eventos ou decisões que se devem verificar para que o nexso causal ocorra entre as atividades, realizações, resultados e impactos. Provas empíricas que apoiem os pressupostos. Ex: Políticas complementares, fatores operacionais, resposta dos atores disponibilidade de serviços que permitem que as realizações produzam os resultados esperados.</p>
INTERVENÇÃO ↓	 <p>A intervenção pretende responder a problemas. Abrange políticas, programas e projetos.</p>
OBJETIVOS ESPECÍFICOS (Relevância) ↓	 <p>O que se pretende com a política em termos práticos.</p>
RECURSOS <i>Inputs</i> ↓	<p>Meios alocados à intervenção para que esta se concretize:</p> <div>  Estratégias  Recursos intangíveis – tempo  Recursos físicos. Por ex. Estrutura Organizacional  Recursos financeiros  Recursos humanos </div>
MECANISMOS	 <p>Os mecanismos são os processos causais que permitem que a intervenção produza resultados. Os mecanismos acionados pelas realizações para produzir os resultados esperados. Correspondem a hipóteses comportamentais e fatores de contexto necessários.</p>
REALIZAÇÕES <i>Outputs</i> ↓	 <p>Atividades das operações, ações. Exemplo: N participantes; N empresas apoiadas; ...</p>
RESULTADOS <i>Outcomes</i> (Eficiência) ↓	 <p>Mudanças previstas que decorrem dos inputs e das realizações da intervenção. Resultados iniciais (mudanças nos conhecimentos, aptidões, capacidades). Resultados intermédios (mudanças comportamentais). Resultados a longo prazo (finalidade das políticas). Os resultados podem ser traduzidos em indicadores de resultado, por exemplo, taxa de pobreza, duração do desemprego.</p>
IMPACTOS* (Eficácia) (Sustentabilidade)	 <p>Mudanças produzidas por uma intervenção, direta ou indiretamente, esperadas ou não esperadas. As evoluções registadas nos indicadores de resultado após a implementação de uma política não correspondem necessariamente ao seu impacto. Os impactos não são imediatamente evidentes. Correspondem aos efeitos da intervenção no contexto socioeconómico. Os impactos devem ser inferidos através de métodos específicos de avaliação, que têm em conta dados e argumentos lógicos.</p>
RISCOS DAS LIGAÇÕES CAUSAIS	 <p>Contexto social, político ou económico que pode afetar a capacidade do programa de alcançar um resultado pretendido. Fatores externos que podem influenciar os resultados esperados. Circunstâncias fora do controlo da intervenção que podem inibir a ocorrência do nexso causal. Elos problemáticos. Por exemplo, desconhecimento prévio dos pressupostos, ou uma investigação anterior que põe em causa os pressupostos.</p>

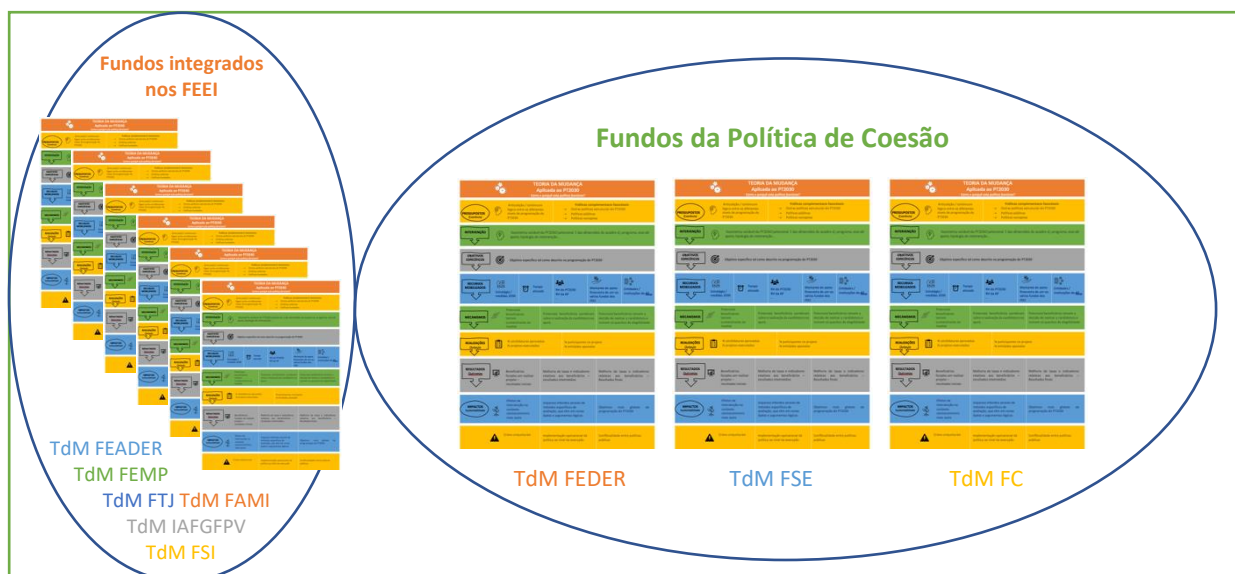
Fonte: Elaboração própria

Tabela 4. TdM aplicada aos FEEI – versão simplificada



Fonte: elaboração própria

Tabela 5. Exemplo de TdMs encaixadas dos FEEI



Notas conclusivas

Após refletirmos sobre os usos da TdM concluímos que – quer na literatura internacional, quer do ponto de vista de alguns atores da comunidade de avaliação dos fundos europeus em Portugal – os desafios da aplicação da TdM aos FEEI, iniciados formalmente desde 2014, têm ainda um relevante caminho a percorrer.

A abordagem metodológica da ABT/TdM, generativa e de contribuição, proposta nos manuais técnicos da Comissão Europeia, e na literatura científica internacional sobre o assunto, e mobilizada na procura de resposta à complexidade da programação dos fundos europeus.

Apesar das robustas soluções teórico-metodológicas apresentadas, e recomendadas pela CE desde 2014, a sua implementação tem representado algumas dificuldades junto comunidade de avaliação portuguesa. Procurámos retratar as perceções de alguns atores de avaliação e identificámos que as suas dificuldades se prendem sobretudo com o cariz abrangente da programação dos fundos, com as dificuldades de operacionalização (falta de capacitação técnica dos RH e de mobilização dos stakeholders), ou ainda com o desfazamento da ambição da envergadura empírica face ao que é exequível na realidade prática.

Após várias recomendações propostas ao longo do texto, propomos uma representação gráfica da TdM-encaixada, enquanto contributo para o caminho de racionalização e simplificação da avaliação baseada na teoria, a percorrer no PT2030.

4. A investigação da ação social e o conceito de mecanismo

Nem sempre admitida, é ambição de qualquer cientista social produzir análises que simultaneamente sejam argutas e desvendem o que não é evidente e indiscutível, porque sustentadas em evidências que não concedem espaço a muitas dúvidas. O desenvolvimento tecnológico das últimas décadas foi bem aproveitado para este propósito da investigação social. A criação e disponibilização alargada de softwares cada vez mais sofisticados de análise de estatística, suportada em robustos sistemas de informação que recolhem e disponibilizam indicadores relevantes e fidedignos de forma regular, permitiu melhorar substancialmente a identificação de dinâmicas de desenvolvimento que afetam os contextos e a forma como se relacionam as mudanças ocorridas e políticas implementadas. As abordagens quantitativas e experimentais ou *quasi* experimentais – por exemplo, as análises contrafactuais – têm permitido, com cada vez maior eficácia, testar os efeitos de intervenções, ultrapassando limitações da formulação inicial, nomeadamente a perspetiva da causalidade única. Apoiados em programas cada vez mais sofisticados, os modelos de análise contrafactual são cada vez mais complexos e surgem agora sob a forma de “structural equations” ou “causal modelling frameworks”.

Porém, se os métodos puramente quantitativos conseguem, com alguma segurança, identificar resultados e atribuí-los a uma intervenção, nomeadamente verificando o que acontece na sua ausência, não conseguem explicar como esse resultado se verificou. Assim, por mais eficazes que seja, estas abordagens permitem responder às questões: “*A intervenção foi eficaz face aos objetivos que prosseguia? Em que medida os resultados observados podem ser diretamente atribuídos à intervenção?*”, mas não permitem responder à questão: “*De que forma (como, porquê e em que contexto) a intervenção produziu (ou não) os efeitos esperados?*”.

Não é ignorado por quem faz investigação em ciências sociais que as explicações de ações sociais são necessariamente complexas. Em análise e avaliação de políticas, mesmo os projetos menos ambiciosos, sustentam-se num conjunto de processos, síncronos e assíncronos, onde atores diversos e com motivações e interesses distintos interagem num ambiente em que o conjunto de oportunidades e custos de cada decisão é afetado por externalidades. Nestas circunstâncias, muitos defendem que explicação causal não poderá ir muito além do enunciado de uma pluralidade de condições determinantes cuja verificação simultânea confere um certo grau de probabilidade à ocorrência de determinados resultados. Na prática, quando se trata de fenómenos sociais e – ainda mais – políticas complexas – é sabido que nenhuma causa leva, em todas as circunstâncias e condições, a um mesmo efeito e, na investigação em ciências sociais e avaliação e políticas, existe já familiaridade com a ideia de configurações causais.

Nos fenómenos sociais a causalidade é eminentemente conjuntural e assimétrica. A causalidade conjuntural implica que os efeitos são frequentemente produzidos por uma combinação de condições, causas múltiplas ou equifinalidades. Implica que várias dessas combinações podem ser suficientes para o mesmo resultado. A assimetria causal refere-se ao facto das causas necessárias e suficientes geralmente só trabalhar num só sentido (Haesebrouck & Thomann, 2021).

Esta perspetiva da causalidade enquanto processo complexo, contingente e intersubjetivo, particularmente ajustada à análise de processos de mudança social tais como os que se realizam na implementação de políticas, sustenta-se na ideia de mecanismo. Na perspetiva clássica (Merton, 1967; Elster, 1989), os mecanismos são construtos teóricos (explicações) que apresentam ligações hipotéticas entre eventos observáveis. Estes mecanismos, fora de uma situação experimental, isto é, em sistemas abertos, são afetados por outros mecanismos que operam ao mesmo tempo: há uma codeterminação causal que pode alterar, e frequentemente altera, os efeitos de um mecanismo considerado isoladamente. É esta codeterminação dos sistemas abertos e que torna a atividade experimental irrealista (Bhaskar, 1997) e exige a utilização de métodos capazes de captar a complexidade. A definição mais genérica, robusta e assente na perspetiva de causalidade surge nas propostas de avaliação realistas e é a de que os mecanismos são “entidades, processos ou estruturas subjacentes que operam em contextos particulares para gerar resultados de interesse” (Pawson & Tilley, 1997; Astbury & Leeuw, 2010).

Os mecanismos geram resultados. Ou seja, são observáveis pelos seus efeitos e é possível reconstruir a trajetória dos seus efeitos a partir de experiências humanas, não redutíveis ao indivíduo. Não são observáveis em cada caso, nem aferíveis a partir da relação entre variáveis na população, mas encontram evidência em ambos os exercícios empíricos, emergindo mais evidentemente em teorias de médio alcance (Merton, 1967). São os processos sociais a partir dos quais as ações individuais e coletivas se constituem como regularidades sociais, passíveis de incorporar uma teoria explicativa. Ligam agência e estrutura. A sua investigação encontra-se num nível de análise intermédia, entre a descrição pura e o *story telling* e a identificação de leis universais gerais (Hedström & Swedberg 1998).

Os atores sociais atribuem sentidos subjetivos – razões - às suas ações, às ações de outros e reagem aos ambientes mais amplos nos quais se encontram. Os mecanismos são as razões e raciocínios que operam a nível individual em conjuntos da população e que explicam o tipo e grau de adesão à proposta de mudança representada numa proposta política. O significado intersubjetivo é assim parte integrante da explicação causal assente em mecanismos.

O interesse na identificação dos mecanismos na análise de fenómenos sociais tem um património alargado nas ciências sociais e, desde a sua génese, também na avaliação de políticas, em especial nas avaliações baseadas na teoria (Chen, 1990; Pawson & Tilley, 1997). Operacionalizado desde logo nas primeiras propostas para uma avaliação baseada na teoria (Chen & Rossi, 1987) assume, na avaliação realista de Pawson & Tilley (1997), um papel de destaque, conceptualizado a partir dos princípios de explicação da causalidade do realismo científico (Baskhar, 1997; Archer, 1995) enquanto “ideias e oportunidades” que promovem a adesão dos grupos ao programa, nas condições e contextos apropriados (Pawson & Tilley, 1997: 57). O mecanismo não é o programa, mas a resposta gerada dos indivíduos e grupos às atividades do programa (Weiss, 1997:46). É descrito em termos de entidades envolvidas em atividades (Beach & Pedersen, 2013). A descrição de um mecanismo num modelo de intervenção implica a análise da combinação entre os recursos (meios e estratégias) disponibilizados no programa e a forma como os stakeholders os interpretam. É a combinação entre mecanismo e contextos – entendidos “fatores e condições, riscos e estratégias de mitigação empregues” - que gera o resultado (Pawson & Tilley, 1997).

Contudo, pese embora se trate de um conceito muito utilizado na linguagem da avaliação, a importância dos mecanismos parece ainda não ser bem compreendida pelos avaliadores (Astbury & Leeuw, 2010:365), que adotam o foco limitado da abordagem a partir de quadros lógicos (LogFrames). A teorização dos mecanismos causais nas abordagens baseadas na teoria é considerada insuficiente, em particular por não identificarem os vários elementos que compõem os mecanismos, constituindo uma base frágil para produzir inferências sobre os mecanismos causais associados a um resultado (Schmitt & Beach, 2015: 431). Ainda que teoricamente os mecanismos assumam um lugar central na AbT, na prática, são operacionalizados como sequências de eventos que ligam dois acontecimentos contingencialmente (como na avaliação realista) ou os elementos constituintes da relação causal são apresentados como suposições não teorizadas (como na análise da contribuição).

A proliferação de conceitos de “mecanismo”, apontada e analisada por vários autores (Mahoney, 2001, 2003; Gerring, 2007; Hedström, 2005) não facilita a clareza e estes surgem operacionalizados em diferentes avaliações como atividades do programa ou enquanto variáveis: independentes, contributivas ou moderadoras, que procuram explicar a correlação existente entre variáveis. Porém, ao contrário das variáveis, os mecanismos não são atributos observáveis de uma unidade de análise, antes procuram explicar a relação entre variáveis. Os mecanismos, nesta perspetiva, residem num nível de abstração acima das variáveis. Uma variável moderadora é, antes, uma demonstração empírica de um mecanismo que se procura identificar e descrever (Astbury & Leeuw, 2010). Operacionalizando o conceito de mecanismo de Robert Merton (1967), Hedström & Swedberg 1998 postulam uma tipologia de mecanismos: macro-micro; micro-micro e micro-macro

Tabela 6. Tipos de Mecanismos

Tipos de Mecanismos	Nível a que opera	Exemplo
Situacionais	macro-micro	Profecias auto-realizáveis; mecanismos de formulação de preferências
Formação da Ação	micro-micro	Dissonância cognitiva; ética protestante e capitalismo
Transformacionais	micro-macro	“pensamento de grupo”; “efeito de manada”

Para explicar os mecanismos situacionais e dar o exemplo de explorações de mecanismos deste tipo – macro-micro - é referenciada a investigação de Erwin Goffman (1963) sobre o comportamento em espaços públicos e a análise situacional de Karl Popper. Estes são mecanismos que ligam a estrutura social e o contexto macro social às crenças, valores, intenções e oportunidades dos sujeitos. As

profecias autorealizáveis ou o efeito pigmaleão⁹, que evidenciam como a categorização social influencia a interação social são outros exemplos de mecanismos deste tipo.

Os mecanismos de formação da ação explicam como uma combinação específica de desejos individuais, crenças e oportunidades de ação geram uma ação específica. É o campo dos mecanismos psicológicos, associado às teorias de formação da decisão, mas também da relação entre a ética protestante e o espírito do capitalismo explorada por Max Weber (2001): as crenças do indivíduo originam uma ação individual específica ao nível económico que potencia a emergência do capitalismo. No caso dos mecanismos transformacionais, os indivíduos interagem e o mecanismo é o processo a partir do qual estas interações individuais se transformam num resultado coletivo, ainda que por vezes não intencional ou esperado. Um outro exemplo dado são os modelos de mercado neoclássicos.

O foco nos mecanismos representa uma postura epistemológica que se ajusta a abordagens metodológicas variadas. Nas abordagens mais quantitativas, uma análise dos mecanismos permite distinguir entre causalidade e correlação, respondendo ao problema das variáveis de confusão, i.e. aquelas que interferem na relação entre duas variáveis de interesse (Steel, 2004). Os métodos contrafactuais procuram testar a presença de mecanismos mediante a derivação lógica dos efeitos que deveriam ser observados se o mecanismo funcionasse como é assumido na teoria e comparando estas assunções com a realidade. Em linguagem quantitativa, um coeficiente de regressão é também uma manifestação de existência de um mecanismo (Hedström & Swedberg 1998).

Numa abordagem mais qualitativa, a causalidade a nível da população é perspectivada como epifenómeno de processos que ocorrem ao níveis e casos (Mahoney, 2008). Estes são, assim, ontologicamente precedentes em relação às causalidades identificadas a nível da população. Se, numa perspectiva da investigação que tem por base uma população, a causalidade é uma probabilidade estatística que opera probabilisticamente nessa população, na ótica de uma investigação com base em estudos de caso, a causalidade é um processo que possibilita ou gera resultados específicos em dadas circunstâncias. Num modelo hipotético-dedutivo de pesquisa, os estudos de caso permitem investigar aspetos das teorias – relações entre variáveis – que foram identificadas a partir de uma análise de variáveis de interesse ou que são assumidas a partir de uma TdM e constitui-se como uma metodologia relevante para a construção de teorias (Ulrichsen & Dadalauri, 2016)

A investigação em torno dos mecanismos sociais define um desenho de pesquisa orientado para a identificação e compreensão das práticas sociais prevalecentes que, quando conectadas, produzem uma relação causal (complexa) que liga uma variável a outra. A Teoria do Programa são as hipóteses de configurações de mecanismos e contextos que se espera darem determinados resultados. Uma das alternativas para a investigação de relações causais a partir de mecanismos que tem revelado interesse dos avaliadores e resultados relevantes é o *Process Tracing*.

⁹ Também conhecido por efeito Rosenthal, aponta como as expectativas dos professores podem afetar o desempenho dos alunos, sendo mais positivos quando as expectativas são positivas e piores face a expectativas negativas.

Process Tracing

Mahoney (2012) designa *Process Tracing* como método das “observações de processos causais”¹⁰ e define como consistindo de generalizações abrangentes que são identificadas em sequências de estudos de caso. Esta Análise de Configurações Causais (ACC) constitui um método interpretativo que articula um raciocínio dedutivo – a partir da teoria - com uma abordagem indutiva, em que a teoria é investigada, afinada e caracterizada nos seus aspetos constituintes.

A vantagem do método é a sua adequação para relacionar teoria (daí a sua aplicação em AbT) e evidência, descrevendo a forma como cada elemento contribuiu para o resultado, permitindo aceder a elementos de explicação não previstos na cadeia causal e que são condições integrantes. Por esta razão a ACC é muito adequada numa avaliação *ex post* baseada na teoria, quando já existem resultados observáveis e uma hipótese acerca da forma como estes foram atingidos (TdM). Contudo, ao contrário das análises contrafactuais, a ACC não exige uma baseline de informação estatística contrafactual e a possibilidade de comparar grupos beneficiários e contrafactuais (Putton & Welle, 2015b).

Geralmente não é possível realizar experiências para testar relações causais em ciências sociais e raras vezes é possível medir o efeito de várias hipotéticas causas contribuidoras para um determinado resultado. É assim muito difícil fazer inferências causais a partir de dados estatísticos sem qualquer conhecimento dos processos sociais subjacentes que permita circunscrever o campo. A ACC, a partir da introdução de protocolos de investigação interpretativa de políticas, permite fornecer informação sobre mecanismos sociais relevantes, constituindo um instrumento útil para compreender e avaliar as mudanças produzidas por políticas e programas (Steel, 2004: 70). A partir de uma lógica Bayesiana¹¹, aditiva e cumulativa a teoria é recorrentemente testada e afinada em testes empíricos, permitindo apreender a complexidade dos fenómenos em estudo.

A ACC procura decompor uma intervenção nas suas diferentes componentes causais – mecanismos hipotéticos - estabelecendo um protocolo de investigação para as analisar e analisar a forma como se relacionam. Esta abordagem é particularmente relevante em situações em que os processos de mudança possuem uma dimensão imaterial e são menos previsíveis, em que múltiplos processos ocorrem em paralelo e em simultâneo e a mudança não pode ser atribuída a uma única causa. É também uma metodologia adequada em intervenções novas ou muito complexas, quando os mecanismos causais são menos conhecidos ou em que existem relações entre vários elementos em interação e se antecipam relações entre eles. A ACC é um método de investigação e análise sistemática particularmente adaptado a contextos complexos em que existem explicações causais concorrentes (Lintelo et al. 2016: 10). A ACC pode ser usada com diferentes propósitos: para testar teorias, para conceber teorias e para explicar resultados (Beach & Pedersen, 2011:2).

¹⁰ Utilizamos neste estudo a tradução “análise de configurações causais”

¹¹ Que se baseia na integração de uma aferição da robustez da evidência (probabilidade) enquanto elemento de análise qualitativa.

O teste de teorias com recurso a um protocolo de investigação de ACC parte da dedução da hipótese acerca do funcionamento do mecanismo causal com base na pesquisa teórica e exploração bibliográfica e investiga se, no caso em análise, o mecanismo está presente. Este é o caso da maioria das aplicações de ACC em avaliação de programas e políticas, em que as hipóteses de configurações de mecanismos e contextos que se espera darem determinados resultados – a Teoria do Programa – são investigadas. Nestes casos, a investigação é desenhada a partir de um corpo teórico-empírico ou de uma teoria da mudança que especula um conjunto de processos e relações entre fatores, condições e resultados. Quando o corpo teórico-empírico permite identificar variáveis interrelacionadas associadas ao resultado que se está a investigar, mas os mecanismos não estão claramente conceptualizados – o que acontece na maioria dos casos - o primeiro passo na aplicação da ACC será a conceptualização de um mecanismo plausível. Este processo pode ser realizado com recurso a entrevistas a atores envolvidos na programação das políticas ou programas, especialistas e *stakeholders*, mas também com base na deteção de relações estatísticas entre indicadores relevantes e que justificam uma investigação aprofundada (Beach & Pedersen, 2011:8).

Quando utilizado para elaborar teorias, a ACC parte da análise de um processo num caso específico (estudo de caso), procurando inferir o mecanismo causal abrangente que se aplica e explica as evidências encontradas. Este tipo de aplicação da ACC representa a perspetiva que esteve na base da formulação das teorias da ação social que, ainda hoje, servem de referência para a construção dos modelos para dedução hipotética.

Os estudos que se focam casos históricos concretos utilizam a ACC com vista a chegar a uma explicação fundamentada acerca das circunstâncias e condições que geraram o resultado que se pretende explicar e caracterizam-se por se sustentar numa reconstituição de um processo do qual se conhece apenas o resultado e não é já possível observar os processos.

A força da ACC reside na sua capacidade de fornecer densidade e substância verificável às explicações teóricas da causalidade na análise de um processo social. Os desenhos de pesquisa em ACC no domínio das políticas públicas, independentemente de serem orientados para o teste ou produção de teoria, combinam um raciocínio indutivo e dedutivo e utilizam diferentes métodos – análise e tratamento estatístico de bases de dados, inquéritos por questionário, entrevistas, histórias de vida e observação participante e análise de dados administrativos. Incorporam também uma dimensão de análise interpretativa de políticas, focando os contextos e como são percecionadas e recebidas as mudanças, procurando uma descrição substantiva e fundamentada (Fischer, 2003).

A perspetiva é organizar a pesquisa em função da resposta a um problema, materializado numa questão de partida, utilizando para isso toda a informação e todas as técnicas necessárias e disponíveis para responder à questão: “De que forma (como, porquê e em que contexto) a intervenção produziu (ou não) os efeitos esperados?”. As conclusões apresentam explicações para as evidências recolhidas e identificam os atores e as ações que geraram os efeitos investigados (Beach & Pedersen, 2012)

As pesquisas abertas, que se organizam em função da resposta a um problema ou questão, enfrentam desafios acrescidos no estabelecimento de critérios de circunscrição da pesquisa. A total saturação é raramente alcançável e é sempre possível estabelecer ligações a outras questões, desenvolvimentos

paralelos, etc. O adequado nível e grau de contextualização de um tema é uma decisão de pesquisa e o relevante é que seja explícita, consistente e justificada (Howarth, 2005). A ACC propõe circunscrever o campo de análise a partir do processo de dedução da pergunta de partida num conjunto de possíveis explicações – hipóteses explicativas – como protocolo de método. São estas hipóteses que são investigadas, utilizando todos os métodos disponíveis de forma a recolher evidência que possa permitir verificar ou rejeitar as hipóteses levantadas. A saturação surge da confiança de que a evidência recolhida permite sustentar uma determinada explicação.

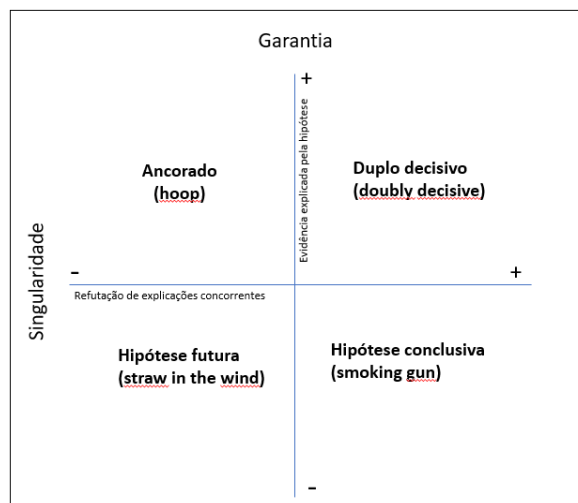
O investigador lista a evidência que, existindo, corrobora a hipótese de trabalho. Podem distinguir-se quatro tipos de evidência de mecanismo (Schmitt & Beach, 2015: 439): a de regularidade, que se relaciona com a existência de padrões estatísticos; de sequências, em que a cronologia temporal e espacial corresponde à teoria explorada; os vestígios, que aponta a existência de um rasto do mecanismo; e a evidência narrativa, relativa a conteúdo do material empírico, como testemunhos, discursos escritos, etc. O contributo original da ACC no âmbito das AbT é o facto da análise da relevância da prova material reunida constituir como uma das etapas do método.

A partir de uma reflexão em torno da relação da evidência recolhida com a hipótese testada – o que poderia designar-se de *prova* - Van Evera (1997:31-2) tipifica em quatro “testes” de verificação da robustez da evidência. São exercícios de análise do corpo empírico da investigação às configurações causais (ACC), refletindo sobre o seu grau de garantia (sustentação) e de singularidade (unicidade) na corroboração ou rejeição das hipóteses de estudo. A evidência é analisada quanto à sua concordância face às hipóteses, ou seja, que parte da evidência recolhida é explicada pela hipótese, e quanto à singularidade da explicação, ou se a evidência permite rejeitar hipóteses concorrentes. Van Evera designou estes testes de doubly decisive (duplo decisivo) hoop test (hipótese ancorada) e Smoking Gun (hipótese conclusiva) e Straw-in-the-wind (hipóteses futuras).

O duplo decisivo (doubly decisive) caracteriza aquelas raras situações em que o tipo de evidência permite corroborar a hipótese e a sua não verificação, rejeitar. É uma hipótese que está bem sustentada em evidência e em que esta que não sustenta qualquer explicação concorrente. Na hipótese ancorada (hoop-test), se a evidência não corroborar a hipótese, pode concluir-se pela sua rejeição, mas a verificação da condição não é suficiente para a sua corroboração (*eg.* é sempre observada nos casos em que o resultado está presente, permitindo eliminar hipóteses concorrentes, mas é também observada em situações em que o resultado não ocorre). A hipótese conclusiva (smoking gun) é aquela em que a evidência corrobora a hipótese de forma muito robusta (difícilmente se observaria, caso a hipótese não se verificasse), mas se a evidência não corroborar a hipótese, não se pode concluir pela rejeição. O teste da hipótese conclusiva aponta condições suficientes para aceitar uma dada hipótese. É um teste que permite inferir causas, especialmente revelante quando os processos não podem ser observados pelo investigador.

A percepção de que, na maioria das situações reais, não existe evidência suficiente para, com confiança elevada, eliminar ou confirmar totalmente hipóteses leva a que, na maioria dos casos, a investigação sobre relações causais em situações de elevada complexidade produza conclusões que assumem a forma de hipóteses futuras (straw-in-the-wind), mapeando a evidência a favor e contra várias hipóteses argumentativas.

Tabela 7. Quadrantes de robustez da hipótese causal



Fonte: adaptado de Putton e Welle (2015b)

Mahoney (2012) aponta quatro questões que devem guiar um exercício de ACC:

- O resultado que se pretende explicar efetivamente ocorreu? Idealmente o resultado esperado está bem definido e conceptualizado à partida e, nesse caso, o desafio é o de demonstrar - e não somente assumir - que o resultado ocorreu no caso em análise. Também os testes de ACC podem ser usados para investigar se existe evidência suficiente de que o resultado ocorreu.
- Os fatores causais investigados podem explicar o resultado? Cada hipótese e cada elemento da cadeia causal deve ser investigado e idealmente alguns testes de análise de ACC devem ser efetuados.
- A causa prevista efetivamente causou o resultado? A resposta a esta questão pode colocar o desafio de reunir evidência de que a causa e o resultado estão presentes, sendo óbvia a relação entre eles ou pode envolver responder à difícil questão de se o resultado é efetivamente consequência do mecanismo identificado.
- Como são eliminadas hipóteses concorrentes? Para muitos resultados, existem múltiplas hipóteses concorrentes. Essas devem ser também investigadas, mas o mais comum é existirem hipóteses alternativas que não podem ser totalmente excluídas, embora possam ser mais inverosímeis, não passando os testes de evidência.

A ACC não se constitui como uma alternativa a outros métodos, mas sobretudo surge como um acrescento relevante à caixa de ferramentas a mobilizar para responder às questões que se colocam em avaliação de programas e políticas. A ACC é próxima da Análise da Contribuição (Mayne, 2001). Partilha com a ACC a perspetiva focada na complexidade causal e nos mecanismos, a base na TdM e o processo dedutivo-indutivo de pesquisa, a multiplicidade de técnicas de investigação e tipos de

evidência analisadas. Mas o foco deste esforço é feito a partir do programa e da análise do impacto da proposta de mudança – programa ou política – e não a explicação de um resultado observado. A questão de partida é se e como funcionou o programa e não – como na ACC - que mecanismos, que processos sociais, espoletados pelo programa ou não, geraram o resultado de interesse. A ACC acrescenta a e reforça a Análise da Contribuição, a partir da definição dos testes à evidência, referidos anteriormente. Priscilla Álamos-Concha et al (2020) utilizam a ACC em articulação com a Qualitative Comparative Analysis (QCA) e Andrew Anguro (2019) sugere que o melhor cenário é quando é possível analisar o impacto de um programa utilizando uma combinação de análise contrafactual, seguida de ACC.

Aplicações de Análise de Configurações Causais

A ACC inscreve-se na tradição de pesquisa compreensiva, constituindo elemento central desta posição epistemológica a densidade e profundidade das investigações conduzidas e que é particularmente ajustada à investigação de mecanismos. A descrição de um mecanismo implica a identificação das condições conjunturais que fazem parte do processo, os seus agentes, perceções e ações (Pawson and Tilley, 1997). No campo da avaliação e fundos comunitários, esta abordagem pode constituir um desafio, já que, para além da complexidade resultante da coexistência de diferentes estratégias e níveis de mudança, os programas comunitários tendem também a ser de grande dimensão e abranger contextos muito diferentes. Em programas complexos coexistem vários mecanismos em ação, cruzando-se efeitos diretos e indiretos, intencionais e não intencionais de múltiplas intervenções. Verificamos, porém, que é justamente em programas complexos que a ACC tem vindo a ser aplicada. O potencial do método para a captação de efeitos sinérgicos e cruzados de programas complexos parece ser suficientemente relevante para aceitar o desafio de analisar, em profundidade, a complexidade. Nestas circunstâncias, a avaliação através da utilização da ACC não deverá estar orientada para a avaliação do programa como um todo, mas antes para a investigação e avaliação de um caso particular, de um resultado, circunscrevendo-o às condições que se espera relevantes estarem presentes para o gerar (Schmitt & Beach, 2015: 434). A partir de um protocolo de ACC, a investigação foca-se apenas num determinado resultado ou efeito e numa determinada ação ou causa. Cada mecanismo é, tanto quanto possível, analisado de forma isolada. Uma análise de programas complexos utilizando a ACC poderá ser realizada num protocolo de investigação etápico que parte de uma análise da coerência, seleccionando as áreas e medidas em que se identificam maiores sinergias para escolher os resultados de interesse que deverão ser analisados. Ou pode constituir uma primeira etapa de investigação num programa de avaliação que, numa segunda fase, é submetido a uma Análise da Contribuição.

O estudo promovido sobre a “Universal Health Care Campaign - UHCC” (Stedman-Bryce 2013) no Gana em 2012-13 é um exemplo citado de protocolo de investigação que utiliza a ACC em combinação com a Análise de Contribuição para avaliar os impactos de uma intervenção (Puton & Welle, 2015). Este programa promovido pela Oxfam tinha como objetivo promover o acesso universal e gratuito à saúde, baseando-se num modelo de governança complexo, uma parceria envolvendo vários tipos de atores – institucionais e líderes locais, públicos e privados, nacionais e estrangeiros - e de áreas diversas. O objetivo foi, entre outros, o de promover o acesso universal e gratuito à saúde, a partir de um conjunto de medidas concertadas, de *lobbying*, divulgação, promoção da investigação e mobilização popular que visavam questionar o modelo de seguro universal adotado e estimular o desenvolvimento de uma

plataforma de apoio ao retorno a um serviço nacional saúde. O desenho de avaliação consistiu no teste deste número limitado de resultados do programa de elevada complexidade, mobilizando testemunhos de informantes privilegiados, documentos administrativos e informação pública. A publicação nas plataformas e redes sociais, por exemplo, constituíram uma fonte de informação e repositório de evidência relevante. Uma das etapas destacadas no relatório foi a reconstrução da TdM do Programa e clarificação de objetivos intermédios, a partir de um workshop participativo com técnicos do Programa, revisitando o modelo lógico (ações, atividades, resultados e impactos). Nesse processo foi definido um objetivo de longo prazo (“melhorar a coordenação das organizações da sociedade civil que defendem o acesso à saúde gratuito e universal”), um resultado de longo prazo (“apoio político para o acesso universal à saúde nas eleições de 2012”) e um de médio prazo (“aumentar o reconhecimento no Gana que o seu atual sistema de financiamento da saúde – seguro universal de saúde - é ineficaz”) para realizar a ACC. Estes resultados foram selecionados de entre os definidos no Programa em função da relevância, por constituírem as áreas de maior dinamismo naquele período e acordados no workshop participativo conduzido (Stedman-Bryce 2013: 14). Cada um dos resultados foi validado e foram identificadas as “narrativas causais”, os mecanismos e componentes que se articulam para o produzir. Identificam-se os agentes e ações, analisando a evidência que suporta ou descarta a hipótese de relação e interação entre atores e eventos. Quase como síntese deste processo, é realizada uma análise da contribuição, apontando o impacto do Programa para cada um dos resultados analisados.

Uma outra aplicação de um modelo de ACC foi o da avaliação dos impactos do “Hunger and Nutrition Commitment Index”- HANCI –que tinha como objetivo identificar durante a implementação do programa, numa avaliação *ongoing*, as formas a partir das quais o programa pode ter contribuído para a mudança de políticas (Lintelo et al. 2016). O programa possuía uma estrutura complexa, articulando rankings, *scorecards*, relatórios de peritos. Envolvia agências de implementação nacionais e internacionais e abrangia Bangladesh, Malawi, Nepal e Zâmbia. O objetivo era perceber em que medida e de que forma o sistema de indicadores promovia o compromisso político das elites dirigentes para com o fenómeno medido, operacionalizado como variável explicativa. Neste caso, a luta contra a fome. O facto de se tratar de um projeto em implementação e de um resultado de mensurabilidade difícil espoletou a decisão por aplicar uma ACC de forma a expor os mecanismos que subjaziam ao impacto político nacional da aplicação de índices comparativos internacionais. Foram identificadas três partes do mecanismo, cada qual correspondendo a entidades e ações. Foram exploradas duas narrativas causais: a de que o impacto resultaria da produção da métrica e da existência de uma estratégia de comunicação eficaz, por melhorar a aceitação e utilização por parte das elites e dos intervenientes políticos e uma segunda que destacava o papel das parcerias entre produtores do indicador e os utilizadores da sociedade civil no país para apoiar a advocacia com elites políticas, focando quatro países-alvo. Na segunda fase, para cada componente do mecanismo, foram organizadas as “manifestações observáveis” – evidências. Por fim, a evidência recolhida foi analisada a partir dos quatro tipo de testes de ACC. Foi concluído que o índice terá influenciado a forma como o problema da nutrição e da fome foi abordado, a partir dos mecanismos expressos em ambas as narrativas causais investigadas. O relatório elaborado revela robustez nas suas conclusões, apresentando – em apenas 60 páginas – uma sólida revisão da literatura acerca dos efeitos dos índices comparativos internacionais e, em geral, do comparativismo em política e *advocacy*, para sustentar a narrativa causal investigada. A originalidade nesta aplicação de ACC resulta do facto de ser aplicado em múltiplos estudos de caso, abrangendo quatro países e a robustez do exercício prévio de

identificação do tipo de evidência que permite a passagem do teste. São mobilizadas teorias de médio alcance (aprendizagem política; teoria das redes; teoria da mentalidade do rebanho, etc.) para circunscrever os mecanismos e é recolhida uma quantidade relevante e suficiente de evidência— em particular, informação recolhida no âmbito das atividades de monitorização do Programa, como questionários aplicados no contexto da execução de atividades ou workshops de avaliação com parceiros - para corroborar as conclusões, sendo este *corpus* de pesquisa discutido na sua validade, a partir dos testes de ACC. São realizados cerca de 22 testes à evidência, testes ancorados (hoop), validando condições necessárias, mas não suficientes. Um dos contributos mais interessantes e que revela o aspeto indutivo do método é a verificação da relevância da utilização dos rankings para a contestação política e das batalhas sobre a sua validade com consequências na credibilidade, assim identificando um potencial risco da estratégia da HANCI (Lintelo et al. 2016).

A ACC é também passível de ser utilizada em articulação com a Qualitative Comparative Analysis (QCA), também permitindo análises que combinam abrangência e profundidade. A partir de um desenho multimétodo de QCA e ACC, as condições que permitem a transferência do conhecimento adquirido na formação para as empresas em Antuérpia foram investigadas (Alamos-Concha et. al). Foram selecionados 50 casos – 15 bem-sucedidos e 35 malsucedidos para realizar uma QCA. Numa segunda fase, quatro casos foram aprofundados a partir de um protocolo de ACC, de forma a explorar a forma como os diferentes mecanismos operaram nos resultados. Os resultados revelaram combinações de oito condições suficientes para o sucesso, algumas sendo condições causais e contextuais. Nesta investigação foi escolhido um processo inicial dedutivo de rastreio teórico - porque a análise QCA permitiu identificar os fatores chave (ou combinações de) que conduziam ao resultado - reconstruído com base na literatura e no trabalho de campo. A evidência foi recolhida a partir de entrevistas individuais semiestruturadas com 8 participantes que transferiram com sucesso a sua formação, através de entrevistas telefónicas com todos os gestores de RH e outras evidências, tais como e-mails, atas e relatórios de dois dos casos selecionados para o rastreio do processo. A análise de dados empíricos foi orientada por três exercícios: o que a informação revela; a confiança na fonte; e explicações alternativas. Com informação recolhida a partir de diversas fontes, é possível refletir acerca do que o conteúdo revela em relação à informação esperada e comparar, permitindo aferir a confiança na fonte e a hipótese de explicações alternativas. Foram identificados quatro possíveis mecanismos: sinalização e retenção desencadeada pelo apoio do supervisor; intervenção reforçada de transferência de formação desencadeada pelo apoio de pares; intervenção de autogestão desencadeada pela prevenção; estabelecimento de objetivos e agência do aprendente desencadeada pelo sentido de urgência. Os mecanismos identificados remetem para a sinalização e retenção, transferência de formação reforçada, intervenção de autogestão e agência de aprendizagem. Estes quatro mecanismos atuam nas fases de aprendizagem e desempenho da formação. A sua conceptualização é concebida como um mecanismo complexo (como um sistema) que diz respeito à transmissão dinâmica das forças causais que podem produzir o resultado (eficácia da transferência da formação). A partir deste exercício os investigadores chegam a um conjunto de recomendações relativas aos fatores contextuais relevantes para facilitar a transferência, relativas ao desenho da formação e às condições individuais do formando/trabalhador

5. Operacionalização da Análise de Configurações Causais (Process Tracing)

Partindo do enquadramento teórico efetuado no ponto anterior, o objetivo deste ponto é explicitar a forma como se pode operacionalizar a ACC, procurando numa primeira fase apresentar um roteiro prático de aplicação desta abordagem, e numa segunda fase, aplicar esse roteiro no âmbito de uma Teoria da Mudança (TdM), de forma a ilustrar uma possível forma de utilização em conjunto destas duas abordagens.

Uma sistematização da forma como a ACC pode ser aplicada no âmbito de uma investigação é efetuada no artigo do CDI (2015), que a partir do trabalho de Beach e Pedersen (2013), definem cinco passos chave que serão explorados de seguida:

- Passo 1. Desenvolver um mecanismo causal hipotético
- Passo 2. Operacionalizar o mecanismo causal
- Passo 3. Recolher as evidências
- Passo 4. Avaliar o peso inferencial das evidências
- Passo 5. Retirar as conclusões do exercício

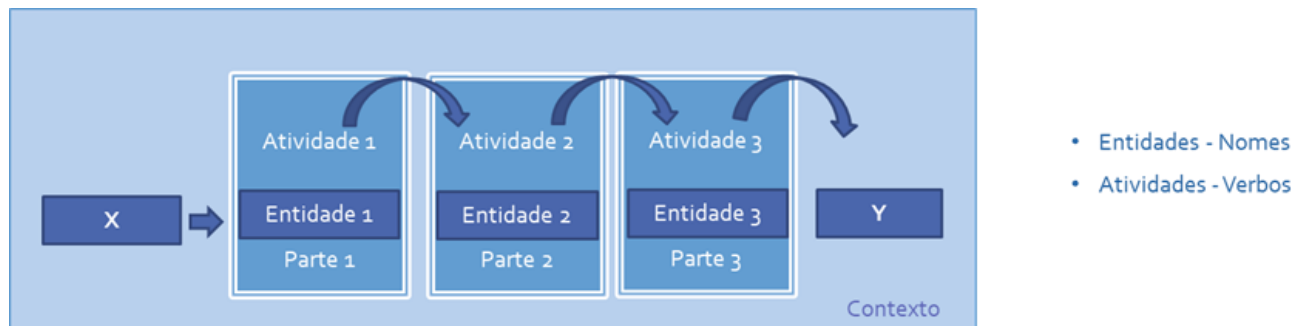
O **primeiro passo** definido é o desenvolvimento do mecanismo causal hipotético, pelo que é importante especificar um pouco mais o conceito de mecanismo e o que envolve a sua construção no âmbito da ACC.

No capítulo anterior ficou explícito que existem múltiplas definições de mecanismos, dependendo do enfoque teórico em que estas são construídas. Para este ponto considerou-se relevante partir de uma definição com um carácter mais operacional, pelo que se terá como base a definição de Glennan (1996) em que mecanismo é um sistema complexo por via do qual se chega a um resultado através da interação de um determinado número de partes. Nesta conceção, um mecanismo pode ser comparado a uma máquina em que cada uma das suas partes pode ser vista como uma roda dentada que transmite energia causal para outra roda dentada (outra parte) contribuindo em última instância para um determinado resultado (Beach e Pedersen, 2013).

Cada uma das partes de um mecanismo causal pode ser conceptualizada como sendo constituída por entidades (pessoas, grupos, estados, fenómenos estruturais, etc.) que levam a cabo atividades que são as produtoras de mudança, são o que transmite as forças causais ao longo do mecanismo.

No momento de construção de um mecanismo, é necessário que contexto onde o mecanismo é ativado seja detalhado, bem como todos os conceitos teóricos devem ser claramente definidos de forma a que todas as partes do mecanismo estejam incluídas e claramente identificadas (Beach e Pedersen, 2013). Numa forma esquemática, mecanismos pode ser representado como na figura abaixo.

Figura 3. Representação esquemática de um mecanismo causal



Fonte: Adaptado de CDI (2015)

Mecanismos são assim considerados teorias sobre como e porquê um evento leva a outro, sendo para tal necessário abrir a “caixa negra da causalidade”, ou seja, apresentar as forças causais (expressa na interação entre as partes) que levam a que a causa X dê origem ao resultado Y.

Tal como já foi referido no capítulo anterior, a elaboração do mecanismo envolve a revisão da teoria (ou TdM), entrevistas com stakeholders, entre outros, de forma a que todos os passos entre X e Y estejam claramente elaborados e que todas as partes do mecanismo possam ser enquadradas enquanto hipóteses a serem testadas (CDI, 2015).

No processo de construção de um mecanismo deve-se ter em atenção que (CDI, 2015; Beach e Pedersen, 2013):

- O mecanismo deve ser dividido no menor número possível de partes em que cada parte causa a parte seguinte (não devem existir “saltos de fé”).
- Cada parte do mecanismo deve ser definida como um elemento individual e necessário de um todo:
 - Cada uma das partes do mecanismo sejam insuficientes, mas necessárias para o mecanismo global;
 - As partes devem ser essenciais para o mecanismo funcionar de tal forma que a ausência de uma das partes faça com que o mecanismo por si só não funcione;
 - Se não se consegue teoricamente argumentar sobre a necessidade de uma parte do mecanismo, ou se posteriormente a análise empírica confirma que uma das partes não é necessária, esta deve ser eliminada do mecanismo para se evitarem redundâncias
- Deve ser possível medir empiricamente cada uma das partes do mecanismo;
- O mecanismo deve ser conceptualizado de uma forma relativamente abstrata face ao caso em estudo para possibilitar a sua utilização noutros contextos.

O **segundo passo** é a operacionalização do mecanismo causal o que implica especificar que evidências empíricas é necessário/possível recolher para determinar se cada parte do mecanismo está presente e se esta tem um link causal com a parte seguinte, sendo esta a base do desenvolvimento das técnicas e instrumentos de recolha de informação. As diferentes partes de um mecanismo podem ter diferentes tipos de evidências relevantes, pelo que estas não são comparáveis (Beach e Pedersen, 2013).

A operacionalização dos mecanismos pode ser sistematizada numa tabela como a que é apresentada de seguida.

Tabela 8. Operacionalização do mecanismo causal

Conceptualização das partes do mecanismo	Evidência prevista de suporte da hipótese	Tipo de evidência utilizada para medir a predição
Parte 1	<i>É expectável que se encontre evidência de que....</i>	<i>Indicar que tipo de evidências serão utilizadas entre os quatro considerados em ACC (regularidade, sequência, vestígios e evidência narrativa). Ex: documentos a consultar, entrevistas a realizar, focus group, estudos de caso, etc.</i>
Parte 2	<i>É expectável que se encontre evidência de que....</i>	<i>Indicar que tipo de evidências serão utilizadas entre os quatro considerados em ACC (regularidade, sequência, vestígios e evidência narrativa). Ex: documentos a consultar, entrevistas a realizar, focus group, estudos de caso, etc.</i>
.....		

Fonte: adaptado de Beach e Pedersen (2013: 112)

Este passo é estabelecido tendo como base a lógica Bayesiana em que é necessário efetuar previsões sobre que evidências devem ser esperadas para ver se a parte de um mecanismo existe e o que conta como evidência para a hipótese alternativa. Desta forma, deve-se igualmente definir explicações alternativas plausíveis- mecanismos/hipóteses alternativas – e procurar as suas evidências de forma a aferir igualmente o seu peso inferencial. No entanto, é frequente não existirem hipóteses alternativas de um mecanismo causal bem como quando estas existem os testes a serem efetuados não podem privilegiar o mecanismo face à hipótese alternativa (Beach e Pedersen, 2013).

Após a operacionalização dos mecanismos e das explicações alternativas, passa-se para o **terceiro passo** que é a recolha das evidências. Devem ser recolhidas evidências para cada manifestação observável de cada parte do mecanismo, em que o avaliador deve avaliar a credibilidade das suas fontes, as suas limitações e enviesamentos (CDI, 2015).

O **quarto passo** é a avaliação do peso inferencial das evidências, seguindo a lógica Bayesiana. Esta em ACC é operacionalizada com a aplicação dos quatro testes já explicitados: *straw-in-the wind tests* (hipóteses futuras), *hoop test* (hipótese ancorada), *smoking gun test* (hipótese conclusiva) e *double decisive test* (duplo decisivo). A construção de testes exige um elevado conhecimento do contexto onde o mecanismo é ativado.

Neste passo o objetivo é sempre maximizar os níveis de certeza (que evidência se deve ter caso a parte do mecanismo esteja presente?) e de singularidade (pode-se explicar a evidência utilizando hipóteses

alternativas?), no entanto quando é necessário tomar uma opção Beach e Petersen (2013) referem que deve ser dada prioridade à certeza, pois cada parte de um mecanismo deve ser vista como necessária. Quanto mais fortes forem os testes mais se pode atualizar o grau de confiança na presença/ausência de um mecanismo, pelo que as inferências sobre a presença de um mecanismo são tão fortes quanto o elo mais fraco dos testes empíricos efetuados para cada uma das suas partes.

Este é um processo com algum nível de subjetividade, uma vez que diferentes investigadores podem atribuir diferentes testes às mesmas evidências, pelo que é fundamental estar claramente especificado as fontes, a natureza das evidências e os testes utilizados (CDI, 2015).

De forma a sistematizar esta avaliação este passo propõe-se a tabela seguinte.

Tabela 9. Proposta de sistematização de aplicação dos testes

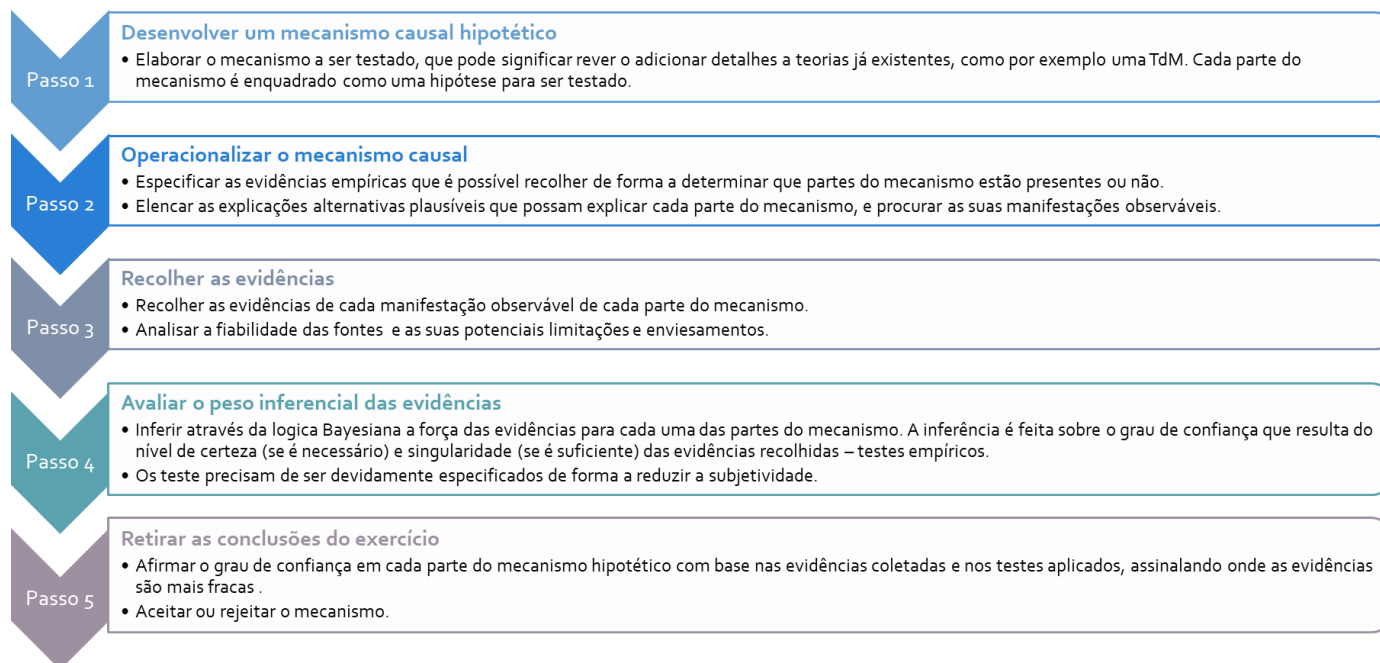
Parte do mecanismo	<i>Parte 1</i>
Evidência obtida	<i>De acordo com as evidências obtidas que indicam que</i>
Especificação do teste	<i>Straw-in-the wind tests (hipóteses futuras), hoop test (hipótese ancorada), smoking gun test (hipótese conclusiva) e doble decisive test (duplo decisivo)</i>
Argumentação	<i>Análise com base nos níveis de certeza e singularidade</i>
Parte do mecanismo	<i>Parte 2</i>
Evidência obtida	<i>De acordo com as evidências obtidas que indicam que</i>
Especificação do teste	<i>Straw-in-the wind tests (hipóteses futuras), hoop test (hipótese ancorada), smoking gun test (hipótese conclusiva) e doble decisive test (duplo decisivo)</i>
Argumentação	<i>Análise com base nos níveis de certeza e singularidade</i>

Fonte: Elaboração própria

Por último o **quinto e último passo** são as conclusões do exercício, em que, com um determinado nível de confiança, se determina a presença de cada uma das partes do mecanismo com base nas evidências recolhidas e nos testes efetuados. O objetivo último é de aceitar ou rejeitar o mecanismo como um todo, sendo que quando este é aceite significa que existem evidências de como e porquê a causa X levou ao efeito Y (CDI, 2015). No entanto se são encontradas evidências fortes que não confirmam uma das partes do mecanismo, não se confirma a existência do mecanismo causal hipotético como um todo. Quando uma parte de um mecanismo não é confirmada, significa que não existe evidência suficiente para confirmar a causalidade, e neste caso descarta-se o mecanismo ou vai-se proceder a revisões teóricas utilizando outras ferramentas na tentativa de detetar um mecanismo/partes de um mecanismo subjacente (Beach e Pedersen, 2013).

A figura seguinte apresenta de forma resumida todos os passos apresentados.

Figura 4. Sistematização dos passos de ACC



Fonte: Elaboração própria a partir de CDI (2015)

Apesar destas partes constituírem-se como um importante roteiro para a aplicação da ACC, é importante ressaltar que estes podem ser de difícil aplicação na utilização da ACC numa avaliação de impacto (CDI, 2015; CDI, 2015a):

- A ACC pode exigir um consumo elevado de tempo uma vez que pode ser necessário uma consulta intensiva a stakeholders e utilização de teorias secundárias, não só para testar os mecanismos hipotéticos definidos, como também as explicações alternativas (hipóteses alternativas);
- A recolha de informação correta para a realização dos testes requer um conhecimento aprofundado do projeto, tempo e capacidade para analisar toda a informação, pois se as evidências recolhidas não forem suficientes e adequadas os resultados dos testes podem ser inconclusivos.
- A ACC é de difícil aplicação quando o resultado da intervenção não é conhecido até ao fim da avaliação ou quando existem múltiplos resultados.

Como é referido por Beach e Pedersen (2013) em grande parte dos estudos é difícil aplicar a ACC para a conceptualização e operacionalização de mecanismos causais. É necessário fazer a sua adaptação e aceitar todas as limitações que a conceptualização e operacionalização dos testes empíricos possam ter, +procurando sempre que estes se aproximam o máximo do ideal.

De forma a se obter uma ideia mais prática da forma de aplicação da ACC, de seguida será efetuado um exercício que pretende ser uma tentativa de aplicar esta metodologia em conjugação com a TdM. Para tal vai ser utilizada a TdM desenvolvida no âmbito da Avaliação Intercalar do Programa Operacional Inclusão Social e Emprego para as tipologias de operação 1.01 Estágios para adultos e 1.02 Apoios à contratação.

4.1 Teoria da Mudança e Análise de Configurações Causais – o caso prático

O objetivo deste exercício não é refazer todo o processo avaliativo aplicando a ACC, mas sim utilizar a TdM, os mecanismos e os resultados obtidos no âmbito da avaliação das tipologias de operação em causa e utilizá-los e adaptá-los, sempre que necessário, de forma a procurar aplicar os passos elencados para a realização da ACC. Neste sentido, e uma vez que não fazia parte da metodologia do exercício de avaliação em causa a aplicação deste método, a utilização deste exemplo apresenta algumas limitações, como se poderá ver de seguida, sendo estas sempre assinaladas e efetuadas propostas de melhor adequação à utilização da ACC.

Os estágios para adultos e os apoios à contratação – enquadramento da teoria da mudança

No âmbito da Avaliação Intercalar do Programa Operacional Inclusão Social e Emprego (PO ISE, 2022a), entre outras metodologias utilizadas para avaliar várias tipologias de operação do programa, foi utilizada a Teoria da Mudança (TdM) e a análise de impacto contrafactual para avaliar os impactos das tipologias de operação 1.01 Estágios para Adultos e 1.02 Apoios à contratação, inseridas no OE1.1 “Integrar, de forma sustentada, adultos e jovens, desempregados ou inativos no mercado de trabalho”.

De acordo com o texto do PO citado no Relatório Final da avaliação (PO ISE, 2022) está incluído na tipologia Estágios:

“(...) ações que visam criar condições adequadas à promoção do emprego de desempregados, apoiando a transição entre o sistema de qualificações e o mercado de trabalho e a integração profissional de desempregados com particulares dificuldades no acesso ao mercado, através da realização de uma experiência prática em contexto de trabalho em entidades empregadoras. Saliente-se que Portugal segue a recomendação do Conselho Europeu da UE relativa ao quadro de qualidade para os estágios e valoriza, na aprovação das candidaturas, critérios de empregabilidade, nomeadamente impedindo o recurso sucessivo a esta medida por entidades que não contratem uma determinada percentagem de estagiários no final” (pág. 105)

Ao passo que os Apoios à contratação são:

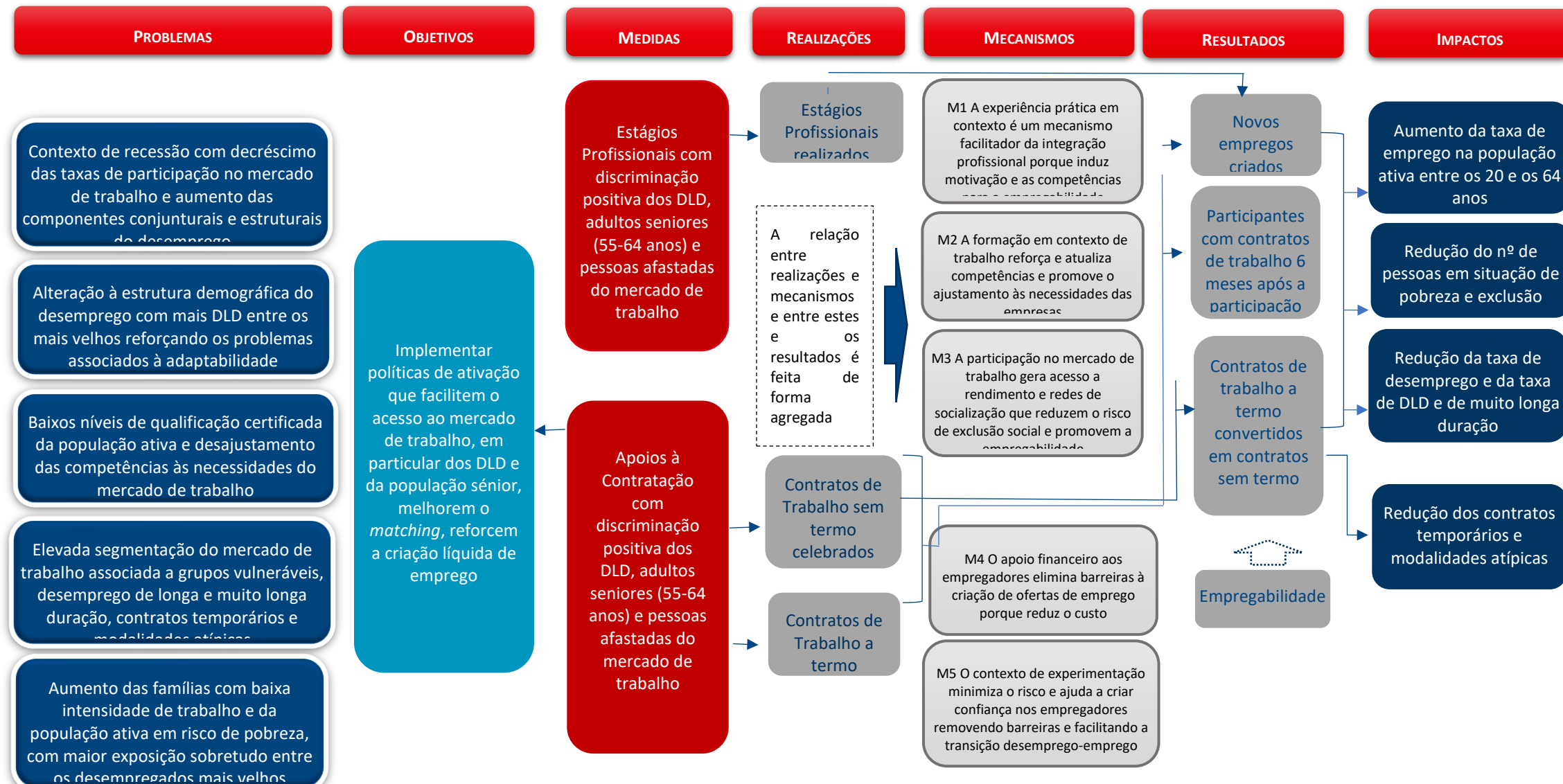
“(...) um conjunto de ações que visam estimular a criação líquida de emprego, mediante o apoio à celebração de contratos de trabalho, no sentido de promover a integração profissional de desempregados. É normalmente majorado o apoio a desempregados mais afastados do mercado de trabalho, tais como: desempregados de longa duração (inscritos há pelo menos 12 meses consecutivos); beneficiários do Rendimento Social de Inserção; pessoas com deficiência e incapacidade; com idade igual ou superior a 45 anos; vítimas de violência doméstica; desempregados cujo cônjuge também está em situação de desemprego, entre outras.” (pág. 105)

Apesar destas medidas incluírem desempregados jovens, uma vez que abrangem desempregados maiores de 18 anos no caso dos estágios e com idade legal para trabalhar no caso dos apoios à contratação, no âmbito da avaliação o público-alvo constitui-se como os desempregados adultos (com idade superior a 29 anos). De acordo com o referido pela equipa de avaliação (PO ISE, 2022) num contexto recessivo do mercado de trabalho (como o que se verificava no momento da programação) o perfil dos desempregados mais velhos com baixas e desadequadas qualificações promove a exposição ao desemprego de longa duração e consequentemente torna-os mais expostos ao risco de pobreza. Este é o contexto de partida que suporta a lógica da programação destas tipologias a maiores de 29 anos pelo que torna necessário apoiar a transição dos adultos desempregados para o mercado de trabalho, através da valorização da estratégia de reforço da adequação de competências e qualificações tendo como prioridade os desempregados em situação de maior vulnerabilidade.

Estas duas tipologias de operação, na programação surgem como um conjunto de respostas que partilham o mesmo racional em termos de problemas e estratégia de ação possuindo um modelo de intervenção onde não são adotadas estratégias diferenciadas em função do perfil dos candidatos, o que permitiu a sua análise integrada e a definição de mecanismos comuns.

De uma forma esquemática a TdM foi explicitada da seguinte forma:

Figura 5. Esquema da Teoria da Mudança das TO 1.01 – Estágios para adultos e 1.02 – Apoios à contratação de adultos



Pressupostos

O contexto de trabalho é potenciado como contexto de aprendizagem.

A função tutoria assegura o desenvolvimento de competências.

Os critérios de seleção dos desempregados e dos beneficiários são ajustados, asseguram o *matching* do perfil às necessidades da empresa e previnem o efeito de substituição.

Riscos

Evolução recessiva do contexto económico

O perfil de alguns desempregados abrangidos (afastados há muito tempo do mercado de trabalho e com competências muito desajustadas) não se adequa ao tipo de intervenção, justificando, por exemplo, ações preparatórias de orientação e desenvolvimento pessoal prévias à inserção do mercado de trabalho.

Encaminhamento para medidas em função da disponibilidade da oferta e não dos perfis dos desempregados.

Prevalência de atitudes de etarismo e estigma em relação aos grupos com maiores dificuldades de integração (desempregados seniores), por parte dos empregadores.

Capacidade de absorção do mercado de trabalho dos postos de trabalho criados em estágios.

Nivelação das remunerações no mercado de trabalho pelo valor do estágio.

Passo 1. Desenvolvimento de mecanismos causais

Num exercício de ACC estuda-se um caso único e analisa-se apenas um mecanismo composto por várias partes e é cada uma dessas partes que vai ser testada. No entanto, ao se conjugar estas duas abordagens – TdM e ACC – devido à aplicação da TdM a programas e intervenções complexas, e igualmente como já referido no capítulo anterior, esta tendencialmente engloba não um mecanismo, mas vários mecanismos causais que procuram especificar a forma como a implementação daquele programa/medida leva a determinados resultados.

Como se pode observar no esquema anterior, foram enunciados pela equipa de avaliação cinco mecanismos. É referido na explicitação da TdM (PO ISE, 2022) que nos estágios e apoios à contratação é valorizado o contexto de exercício de uma atividade como espaço de aprendizagem que por sua vez permite fazer face ao problema da obsolescência das qualificações de alguns grupos de adultos desempregados. Nos estágios a aprendizagem é mais intencional e estruturada ao passo que nos apoios à contratação esta é dinâmica e implícita ao contexto de exercício da atividade profissional. Neste contexto a equipa definiu dois mecanismos:

M1: A experiência prática em contexto é um mecanismo facilitador da integração profissional porque introduz motivação e as competências para a empregabilidade.

M2: A formação em contexto de trabalho reforça e atualiza competências e promove o ajustamento às necessidades das empresas.

Por outro lado, a participação no mercado de trabalho permite a obtenção de rendimentos e de redes de socialização, pelo que a equipa de avaliação refere que as medidas em avaliação contribuem para mitigar os problemas de exclusão e o risco de pobreza que afetam os desempregados de longa duração. Foi então definido como terceiro mecanismo da avaliação:

M3: A participação no mercado de trabalho gera acesso a rendimentos e redes de socialização que reduzem o risco de exclusão social e promovem empregabilidade.

Por último, de acordo com a equipa de avaliação, nestas medidas o apoio financeiro é uma forma de serem viabilizadas novas ofertas de emprego e de estágio e desta forma aumentar a disponibilidade dos empregadores para recorrerem a desempregados do serviço público de emprego e assim avaliarem as suas capacidades para desenvolver uma atividade profissional específica, estando mais uma vez implícita a aquisição de competências alinhadas com o posto de trabalho. No caso dos estágios o financiamento é relevante para possibilitar estratégias de aprendizagem mais estruturadas e exigentes, ao passo que nos apoios à contratação o incentivo financeiro é um apoio direto à adaptação do posto de trabalho. Neste contexto a equipa de avaliação definiu dois mecanismos.

M4: O apoio financeiro aos empregadores elimina barreiras à criação de ofertas de emprego porque reduz o custo.

M5: O contexto de experimentação minimiza o risco e ajuda a criar a confiança nos empregadores removendo barreiras e facilitando a transição emprego-desemprego.

É importante referir que nesta TdM, e de acordo com a equipa de avaliação, “a forma como é explicitado o quadro lógico não favorece a identificação de cadeias de relação individualizadas entre realizações, mecanismos e resultados” (PO ISE, 2022: 109), mas sim os mecanismos encontram-se no processo de mediação entre as realização e resultados identificados. Neste âmbito pode-se afirmar que os mecanismos explicitam “resultados intermédios” que em conjunto contribuem para os “resultados finais” das tipologias de operação em análise e que, por sua vez, foram especificados na programação.

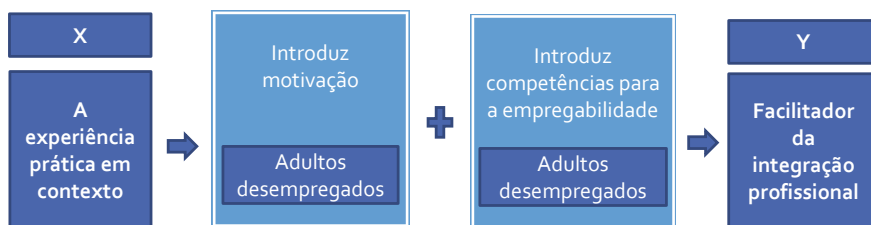
Ainda assim, no âmbito da aplicação da ACC com a TDM é relevante que exista esta relação concreta entre realizações mecanismos e resultados em que os mecanismos explicitam a relação causal através do qual as realizações levam aos resultados, de forma a que a ACC contribuía direta e objetivamente para a validação da TDM.

Passando para o desenvolvimento dos mecanismos, e uma vez que se trata de uma TdM relativamente circunscrita, centrada em duas tipologias de operação específicas, com mecanismos com um número reduzido de partes que levam a resultados semelhantes, considerou-se relevante aplicar a ACC a todos os mecanismos elaborados pela equipa de avaliação e não selecionar apenas um caso particular, pois se tal se fizesse seria redutor para o exercício que aqui é proposto.

Tendo em atenção a definição e representação gráfica de mecanismo causal apresentada no ponto anterior, procurou-se traduzir esquematicamente cada um dos mecanismos de forma a ilustrar o processo de construção de um mecanismo causal. Para tornar o exercício mais claro, foram introduzidos alguns elementos nos mecanismos definidos (como por exemplo a especificação das entidades) de forma a ter um melhor enquadramento no esquema. Por outro lado, tendo em atenção a formulação original dos mecanismos, os esquemas foram tendo igualmente algumas adaptações de mecanismo para mecanismo, uma vez que a representação gráfica trata-se de um ideal tipo.

Assim, o **primeiro mecanismo** “A experiência prática em contexto é um mecanismo facilitador da integração profissional porque introduz motivação e as competências para a empregabilidade” poderia esquematicamente ser representado como na figura abaixo. A forma como o mecanismo está formulado, não indicia tanto a transmissão das forças causais entre as partes (no sentido de que a parte 1 leva ao desenvolvimento da parte 2) mas sim é a associação das duas partes que dão origem ao resultado.

Figura 6. Representação esquemática do mecanismo 1

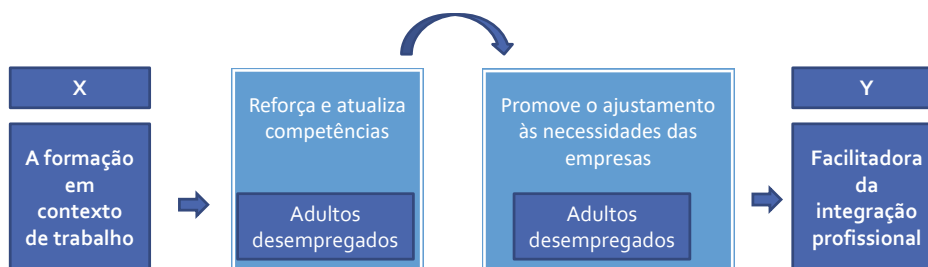


Fonte: elaboração própria

Interpretando o esquema, o mecanismo hipotético pode ser enunciado como: “A experiência prática em contexto de trabalho motiva os adultos desempregados e desenvolve competências para a empregabilidade facilitando a sua integração profissional”.

O **segundo mecanismo** foi definido como: “A formação em contexto de trabalho reforça e atualiza competências e promove o ajustamento às necessidades das empresas”. Nesta formulação não está definido um mecanismo causal pois não está explícita a forma como a formação em contexto de trabalho reforça e atualiza as competências ou promove o ajustamento às necessidades das empresas, sendo que nesta interpretação estes seriam os resultados do mecanismo. No entanto, procurando uma adaptação, pode-se referir que a formação em contexto de trabalho facilita a integração profissional uma vez que permite a atualização de competências e o ajustamento às necessidades das empresas. Desta forma o mecanismo pode ser representado da seguinte forma:

Figura 7. Representação esquemática do mecanismo 2



Fonte: elaboração própria

Assim, o mecanismo causal fica: “A formação em contexto de trabalho reforça e atualiza as competências dos adultos desempregados, promovendo o seu ajustamento às necessidades das empresas o que facilita a sua integração profissional”.

Na figura seguinte apresenta-se a ilustração do **terceiro mecanismo** definido na avaliação como: “A participação no mercado de trabalho gera acesso a rendimentos e redes de socialização que reduzem o risco de exclusão social e promovem empregabilidade”. Este mecanismo indicia dois resultados que se encontram interdependentes - a promoção da empregabilidade e a redução do risco de exclusão social – e que refletem os dois Objetivos Temáticos do PO ISE explícitos no texto da programação:

Objetivo temático 8 – promover a sustentabilidade e qualidade do emprego e apoiar a mobilidade dos trabalhadores;

Objetivo temático 9 – Promover a inclusão social e combater a pobreza e discriminação.

Partindo da formulação de que um mecanismo causal produz um resultado, este mecanismo pode ser desagregado em dois em que os resultados são diferentes:

Figura 8. Desagregação do mecanismo 3

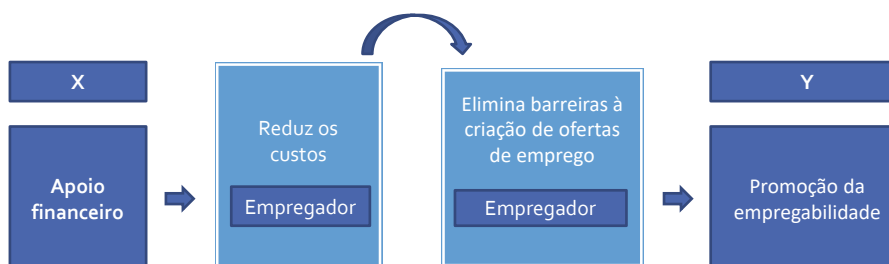


Fonte: elaboração própria

Estes mecanismos ficam então: “A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, promovem a sua empregabilidade”; “A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, reduz o seu risco de exclusão”. Uma vez que na avaliação não se verificou esta desagregação, neste exercício estes mecanismos serão analisados em conjunto.

O **quarto mecanismo** desenvolvido na avaliação, “O apoio financeiro aos empregadores elimina barreiras à criação de ofertas de emprego porque reduz o custo” é um mecanismo que apenas é constituído por uma parte e que tem como resultado a eliminação das barreiras à criação da oferta de emprego. No entanto, a eliminação das barreiras à criação de ofertas de emprego leva, por sua vez, à promoção da empregabilidade, um dos objetivos das medidas que se encontram em análise. Desta forma o mecanismo pode ser esquematizado da seguinte forma:

Figura 9. Representação esquemática do mecanismo 4

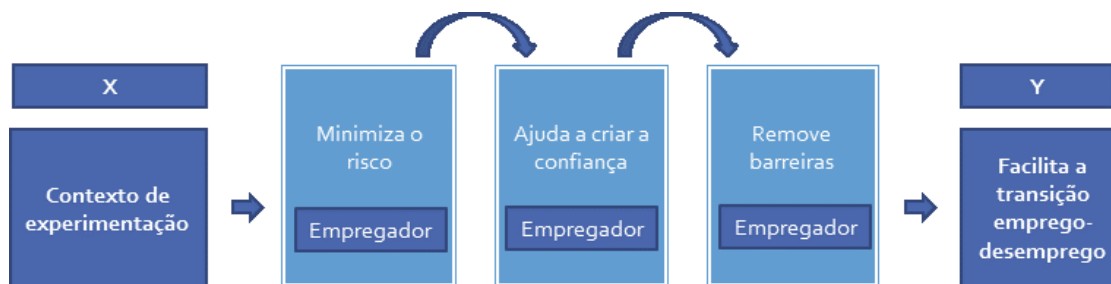


Fonte: elaboração própria

A sua formulação fica: “O apoio financeiro ao reduzir os custos dos empregadores, elimina barreiras à criação de ofertas de emprego o que leva à promoção da empregabilidade”.

Por último, o **quinto mecanismo** foi enunciado como: “O contexto de experimentação minimiza o risco e ajuda a criar a confiança nos empregadores removendo barreiras e facilitando a transição emprego-desemprego”. Esquematicamente fica:

Figura 10. Representação esquemática do mecanismo 5



Fonte: elaboração própria

A formulação fica idêntica à utilizada no âmbito da avaliação: “O contexto de experimentação, ao minimizar o risco dos empregadores, ajuda à criação de confiança o que leva à remoção de barreiras, facilitando a transição do desemprego para o emprego”.

Como se pode observar, os mecanismos apresentados apresentam algumas limitações face ao enunciado no passo 1, uma vez que alguns (como acontece no mecanismo 1 e 3) não apresentam propriamente uma relação causal entre as partes que o constituem sendo mais a associação das partes que leva ao resultado. Por outro lado, a elaboração das partes poderia estar mais detalhada explicando com mais pormenor como uma parte origina a outra.

Passo 2. Operacionalização dos mecanismos causais

Após o desenvolvimento dos mecanismos, passa-se para a sua operacionalização, onde o investigador define que evidência deverá encontrar para medir a predição e que tipo de evidências deverá recolher para cada parte do mecanismo.

Como se está a realizar um exercício a partir de mecanismos que já foram testados no âmbito de uma avaliação, não se irá prever que tipo de evidência será utilizada para medir a predição, mas sim a que foi efetivamente utilizada – resultados da análise de impacto contrafactual. Ainda assim, entre parêntesis colocou-se outros possíveis tipos de evidência que poderiam ser utilizados para medir a predição. Por outro lado, e visto que os mecanismos não são muito complexos no que se refere ao seu número de partes, ter-se-á como base da análise o mecanismo como um todo e não as suas partes separadamente. A tabela seguinte ilustra este passo.

Tabela 10. Operacionalização dos mecanismos causais da TO 1.01 – Estágios para adultos e 1.02 – Apoios à contratação de adultos

Conceptualização do mecanismo	Evidência prevista de suporte da hipótese	Tipo de evidência utilizada para medir a predição
A experiência prática em contexto de trabalho motiva os adultos	Espera-se encontrar evidência junto dos beneficiários das medidas (adultos que frequentaram estágios e apoios à contratação) e junto das entidades empregadoras e acolhedoras de estágio que por terem estado num contexto prático de aprendizagem, ou seja, por	Resultado da análise de impacto contrafactual - evidência de sequência

Conceptualização do mecanismo	Evidência prevista de suporte da hipótese	Tipo de evidência utilizada para medir a predição
desempregados e ao desenvolver competências para a empregabilidade facilita a sua integração profissional	terem tido a possibilidade de conhecer a cultura da empresa e focar a sua aprendizagem numa realidade concreta: - A motivação dos beneficiários das medidas é aumentada; - As competências transversais para a participação no mercado de trabalho dos beneficiários das medidas são desenvolvidas, estando estes mais capacitados para exercer uma atividade profissional.	<i>(Poderiam ser utilizadas entrevistas e focus group a entidades empregadoras e acolhedoras de estágio e beneficiários (evidência narrativa))</i>
A formação em contexto de trabalho reforça e atualiza as competências dos adultos desempregados, promovendo o seu ajustamento às necessidades das empresas o que facilita a sua integração profissional	Espera-se encontrar evidência junto dos beneficiários das medidas e junto das entidades empregadoras e acolhedoras de estágio que por terem estado num contexto prático de aprendizagem, ou seja, por terem tido a possibilidade de conhecer a cultura da empresa e focar a sua aprendizagem numa realidade concreta: - Os beneficiários das medidas atualizaram as suas competências específicas relacionadas com os contextos e as funções que se encontraram a exercer; - A experiência possibilitou um melhor ajustamento das competências dos beneficiários das medidas às necessidades específicas das empresas onde trabalharam.	Resultado da análise de impacto contrafactual - evidência de sequência <i>(Poderiam ser utilizadas entrevistas e focus group a entidades empregadoras e acolhedoras de estágio e beneficiários (evidência narrativa))</i>
A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, promovem a sua empregabilidade	Espera-se encontrar evidência junto dos beneficiários das medidas que estas, ao permitirem a participação no mercado de trabalho; - Aumenta o seu acesso a rendimentos; - Reforçam a sua auto-estima; - Alargam a sua rede contactos e todas as oportunidades que a elas estão associadas.	Resultado da análise de impacto contrafactual - evidência de sequência <i>(Poderiam ser utilizadas entrevistas e focus group a beneficiários (evidência narrativa))</i>
A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, reduz o seu risco de exclusão		
O apoio financeiro ao reduzir os custos dos empregadores, elimina barreiras à criação de ofertas de emprego o que	Espera-se encontrar evidência junto das entidades empregadoras e acolhedoras de estágio que: - Os apoios levam a uma redução efetiva dos seus custos; - A adaptação do posto de trabalho com menos encargos financeiros elimina a barreira financeira o que torna um ponto de	Resultado da análise de impacto contrafactual - evidência de sequência <i>(Poderiam ser utilizadas entrevistas e focus group a entidades empregadoras e</i>

Conceptualização do mecanismo	Evidência prevista de suporte da hipótese	Tipo de evidência utilizada para medir a predição
leva à promoção da empregabilidade	partida para a criação de novos postos de trabalho (incluindo estágios) para desempregados inscritos no serviço público de emprego.	<i>acolhedoras de estágio (evidência narrativa))</i>
O contexto de experimentação, ao minimizar o risco dos empregadores, ajuda à criação de confiança o que leva à remoção de barreiras, facilitando a transição do desemprego para o emprego	<p>Espera-se encontrar evidência junto entidades empregadoras e acolhedoras de estágio que:</p> <ul style="list-style-type: none"> - Estas tipologias de operação reduzem os riscos financeiros da mobilização de novos recursos humanos; - Existe um aumento da confiança dos empregadores nas capacidades dos desempregados inscritos no serviço público de emprego; - Existem menos barreiras o que leva à contratação de desempregados inscritos no serviço público de emprego. 	Resultado da análise de impacto contrafactual - evidência de sequência <i>(Poderiam ser utilizadas entrevistas e focus group a entidades empregadoras e acolhedoras de estágio (evidência narrativa))</i>

Fonte elaboração própria baseado em Beach e Pedersen (2013: 112)

No âmbito desta avaliação, não se identificaram mecanismos alternativos, ainda que alguns dos mecanismos elaborados possam suscitar explicações alternativas para o mesmo resultado, como acontece, por exemplo, com os mecanismos 1 e 2 ou os mecanismos 3 e 4.

Neste caso deveriam ser, por exemplo, definidos mecanismo alternativos para a facilitação da integração profissional de desempregados adultos, para o aumento das ofertas de emprego, etc. procurando analisar posteriormente o seu peso inferencial.

Passo 3. Recolher as evidências

No âmbito da avaliação, foi utilizada a Análise de Impacto Contrafactual aplicada a uma base de dado construída com dados recolhidos no Instituto de Emprego e Formação Profissional (IEFP) e Segurança Social. De acordo com o Relatório Final da avaliação (PO ISE, 2022a) o impacto foi avaliado entre o momento de entrada na intervenção e até 60 meses após, para as intervenções iniciadas no período entre 2014 e 2020.

Foi utilizada uma análise de carácter quase-experimental – analisando os resultados obtidos pelo grupo-alvo da intervenção que por sua vez é constituído por dois (sub)grupos, um referente aos beneficiários dessas medidas (grupo tratado) e outro respeitante aos não-beneficiários (grupo de controlo ou não tratado). O grupo de controlo corresponde a outros utentes do IEFP: a) utentes com mais de 30 anos para a avaliação dos estágios, e utentes com pelo menos 30 anos para avaliação das medidas de apoio à contratação; b) utentes que não beneficiaram de intervenções estruturantes (estágios, contratos de trabalho, etc.); c) utentes que beneficiaram de intervenções de foco pessoal ou de ações de formação, mas não de qualificação, ou seja, exclui formações conducentes a um grau

académico, formações tecnológicas, formações profissionalizantes e afins; d) Intervenções com duração inferior a 90 dias.

O impacto das medidas foi avaliado através de um conjunto de indicadores com base nos dados do IISS: taxa de emprego, remuneração mensal média, taxa de trabalhadores por conta de outrem, e taxa de trabalhadores independentes. De acordo com a equipa de avaliação não foi possível avaliar o resultado nos tipos de contrato a termo ou sem termo (e as suas várias modalidades) uma vez que essa variável não foi fornecida pelo IISS.

Passo 4. Aferir o peso inferencial das evidências

Após a recolha das evidências, vai-se analisar o seu peso inferencial, aplicando os testes a cada um dos mecanismos. Para este exercício, nas evidências obtidas utilizaram-se os resultados da análise de impacto contrafactual que a equipa de avaliação destacou no relatório Final da Avaliação (PO ISE, 2022) como evidências para a análise de cada um dos mecanismos elaborados.

Tabela 11. Análise do peso inferencial das evidências - aplicação dos testes

Mecanismo 1	A experiência prática em contexto de trabalho motiva os adultos desempregados e ao desenvolver competências para a empregabilidade facilita a sua integração profissional
Evidência obtida	<p>Dos resultados da Análise de Impacto Contrafactual, a equipa de avaliação destaca que:</p> <ul style="list-style-type: none"> • A empregabilidade dos beneficiários que não concluíram os estágios (taxa de empregabilidade de 50%), sendo inferior à dos que concluíram os estágios (taxa de empregabilidade de 60%), não deixa de ser alta, refletindo, muito provavelmente, ganhos de participação na medida. A evolução da taxa de emprego é semelhante entre os dois grupos, o que reflete a sustentabilidade dos resultados dos estágios. • Esta perspetiva é reforçada pela inexistência de diferença relevante na remuneração auferida ao longo do tempo entre os que concluíram e os que não concluíram a participação na medida.
Especificação do teste*	<p>À luz das evidências recolhidas na avaliação, é um straw-in-the wind test (hipóteses futuras).</p> <p>Com base nas evidências previstas na tabela anterior, é um smoking gun test (hipótese conclusiva).</p>
Argumentação	<p>Straw-in-the wind tests (hipóteses futuras) – pois as evidências obtidas não são suficientes nem necessárias para confirmar a hipótese, ou seja, a sua obtenção nem valida o mecanismo, nem sugere a validação de eventuais hipóteses alternativas. A evidência apenas sugere que as medidas em análise contribuíram para a empregabilidade dos beneficiários e para os ser valores remuneratórios.</p> <p>Smoking gun test (hipótese conclusiva) - a formulação do mecanismo implica a obtenção de um conjunto de evidências que são relativamente fáceis de se obter e que podem dar uma forte previsão da existência do mecanismo - se os beneficiários e as entidades empregadoras e acolhedoras de estágio, através de entrevistas e <i>focus group</i> apresentarem um discurso no sentido do aumento da motivação e das competências transversais para a participação no mercado de trabalho por parte dos beneficiários. No entanto, se estas evidências não forem obtidas (ausência de evidência) não se possuem elementos suficientes para refutar inequivocamente o mecanismo. Os conceitos elencados no mecanismo como motivação e competências são vastos e podem abarcar explicações alternativas, pelo que podem existir inúmeras razões para as evidências não estarem presentes</p>
Mecanismo 2	A formação em contexto de trabalho reforça e atualiza as competências dos adultos desempregados, promovendo o seu ajustamento às necessidades das empresas o que facilita a sua integração profissional
Evidência obtida	<p>A análise de impacto contrafactual indica que:</p> <ul style="list-style-type: none"> • Entre os beneficiários que concluíram os estágios, a taxa de emprego imediatamente após a conclusão da intervenção é de 60%, ao passo que entre os que não concluíram o valor é de menos 10%. A evolução da taxa de emprego é semelhante entre os dois grupos, o que reflete a sustentabilidade dos resultados dos estágios. • O impacto dos estágios na empregabilidade dos abrangidos pela medida é positivo pois os beneficiários da medida têm uma taxa de empregabilidade que

	<p>corresponde a mais 30 p.p. do que o verificado no grupo daqueles que não participaram na medida.</p> <ul style="list-style-type: none"> • O ganho dos beneficiários da medida estágios em termos de taxa de emprego mantém-se pois ao fim de 60 meses a taxa de emprego entre os que participaram é, ainda, cerca de 20 p.p superior entre aqueles que participaram na medida, correspondendo a uma taxa de 70% para estes e de 50% para os que não participaram na medida. • No caso das medidas de apoio à contratação, os ganhos em termos de empregabilidade são também positivos. A título de exemplo, refira-se que 24 meses após a atribuição do apoio, a diferença da taxa de empregabilidade entre os indivíduos apoiados e os do grupo de controlo se situava em cerca de 40 p.p para todas as operações analisadas (Estímulo 2013, TSU, Estímulo, Contrato Emprego).
Especificação do teste*	<p>À luz das evidências recolhidas na avaliação, é um <i>straw-in-the wind test</i> (hipóteses futuras).</p> <p>Com base nas evidências previstas na tabela anterior, é um <i>smoking gun test</i> (hipótese conclusiva).</p>
Argumentação	<p><i>Straw-in-the wind tests</i> (hipóteses futuras) – argumentação semelhante à do mecanismo anterior pois as evidências obtidas não são suficientes nem necessárias para confirmar a hipótese, ou seja, a sua obtenção nem valida o mecanismo, nem sugere a validação de eventuais hipóteses alternativas.</p> <p><i>Smoking gun test</i> (hipótese conclusiva) – tal como no mecanismo anterior a formulação do mecanismo implica a obtenção de um conjunto de evidências que são relativamente fáceis de se obter e que podem dar uma forte previsão da existência do mecanismo - se os beneficiários e as entidades empregadoras e acolhedoras de estágio, através de entrevistas e <i>focus group</i> apresentarem um discurso no sentido da atualização das competências dos beneficiários e do seu ajustamento às necessidades específicas das empresas onde trabalharam. No entanto, a ausência desta evidência não permite rejeitar inequivocamente o mecanismo, uma vez que tal como o mecanismo anterior a aquisição de competências e o ajustamento às necessidades das empresas poderão ter outras causas, podendo existir hipóteses alternativas para as evidências não se encontrarem.</p>
Mecanismo 3	<p>A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, promovem a sua empregabilidade</p> <p>A participação no mercado de trabalho ao gerar acesso a rendimentos e redes de socialização aos adultos desempregados, reduz o seu risco de exclusão</p>
Evidência obtida	<p>A equipa de avaliação destacou como evidência que:</p> <ul style="list-style-type: none"> • A remuneração média mensal dos adultos abrangidos nos estágios é superior em cerca de 80 euros à saída da intervenção, mantendo-se estável esta diferença ao longo do tempo. • A taxa de trabalhadores por conta de outrem é maior em cerca de 10% (estáveis ao longo do tempo) entre os abrangidos na medida estágios. • Nos apoios à contratação, no Estímulo 2013 a diferença salarial era de cerca de 75 euros no início (tendendo a estabilizar ao longo do tempo para um valor cerca de 7% mais elevado) enquanto na medida Contrato Emprego a diferença se reduz para 30 euros (tendendo a estabilizar em torno dos 6%).

Especificação do teste*	<p>À luz das evidências recolhidas na avaliação, é um <i>straw-in-the wind test</i> (hipóteses futuras).</p> <p>Com base nas evidências previstas na tabela anterior, é igualmente um <i>straw-in-the wind test</i> (hipóteses futuras).</p>
Argumentação	<p><i>Straw-in-the wind test</i> (hipóteses futuras) – por um lado, as evidências obtidas não são suficientes nem necessárias para confirmar a hipótese, ou seja, a sua obtenção nem valida o mecanismo, nem sugere a validação de eventuais hipóteses alternativas (os dados apenas indicam que a participação nas medidas trouxe aos beneficiários melhores remunerações). Por outro lado, face às evidências previstas, e devido à formulação mais abrangente deste mecanismo, a existência de evidências poderia aumentar a plausibilidade da hipótese, mas seria difícil comprová-la inequivocamente uma vez que são múltiplos os fatores que poderão contribuir para o reforço da autoestima e o aumento das redes de socialização dos adultos desempregados, bem como são igualmente múltiplos os fatores que reduzem os riscos de exclusão social. Por outro lado, e pelas mesmas razões, não encontrar evidências também não rejeita a existência do mecanismo, uma vez que a m.</p>
Mecanismo 4	<p>O apoio financeiro ao reduzir os custos dos empregadores, elimina barreiras à criação de ofertas de emprego o que leva à promoção da empregabilidade</p>
Evidência obtida	<p>A análise de impacto contrafactual indica que:</p> <ul style="list-style-type: none"> • Entre os beneficiários que concluíram os estágios, a taxa de emprego imediatamente após a conclusão da intervenção é de 60%, ao passo que entre os que não concluíram o valor é de menos 10%. A evolução da taxa de emprego é semelhante entre os dois grupos, o que reflete a sustentabilidade dos resultados dos estágios. • Os resultados apurados evidenciam, com clareza, que as medidas apoiadas contribuíram positivamente para facilitar o acesso ao mercado de trabalho dos beneficiários da medida estágios e para melhorar as condições de participação • O impacto observado dos apoios à contratação sobre a criação de emprego produziu ganhos limitados em termos do vínculo associado à contratação.
Especificação do teste*	<p>À luz das evidências recolhidas na avaliação, é um <i>straw-in-the wind test</i> (hipóteses futuras).</p> <p>Com base nas evidências previstas na tabela anterior, é um <i>doble decisive test</i> (duplo decisivo)</p>
Argumentação	<p><i>Straw-in-the wind test</i> (hipóteses futuras) – pois as evidências obtidas não são suficientes nem necessárias para confirmar a hipótese, ou seja, a sua obtenção nem valida o mecanismo, nem sugere a validação de eventuais hipóteses alternativas.</p> <p><i>Doble decisive test</i> (duplo decisivo) – a formulação do mecanismo implica a obtenção de um conjunto de evidências que podem dar uma forte previsão da existência do mecanismo, mediante a explicitação por parte dos empregadores que custos os apoios permitem reduzir e que essa redução de custos levou à eliminação de barreiras financeiras e à criação de um determinado número de postos de trabalho/estágios que foram ocupados por desempregados inscritos no serviço público de desemprego. Por outro lado, se não se encontram evidências de redução de custos e criação de novos postos de trabalho/estágios a confiança na existência do mecanismo é claramente reduzida.</p>

Mecanismo 5	O contexto de experimentação, ao minimizar o risco dos empregadores, ajuda à criação de confiança o que leva à remoção de barreiras, facilitando a transição do desemprego para o emprego
Evidência obtida	<p>A equipa de avaliação destacou como evidência que:</p> <ul style="list-style-type: none"> • A remuneração média mensal dos adultos abrangidos nos estágios é superior em cerca de 80 euros à saída da intervenção, mantendo-se estável esta diferença ao longo do tempo. Do mesmo modo, a taxa de trabalhadores por conta de outrem é maior em cerca de 10% (estáveis ao longo do tempo) entre os abrangidos na medida. • Nos apoios à contratação, no Estímulo 2013 a diferença salarial era de cerca de 75 euros no início (tendendo a estabilizar ao longo do tempo para um valor cerca de 7% mais elevado) enquanto na medida Contrato Emprego a diferença se reduz para 30 euros (tendendo a estabilizar em torno dos 6%). Os ganhos observados são também positivos no que diz respeito ao tipo de contrato, sendo superior naqueles que possuem um contrato por conta de outrem entre os abrangidos nos apoios à contratação. • Os dados atestam com consistência que os estágios e os apoios à contratação contribuíram positivamente para o nível salarial dos trabalhadores abrangidos pelas medidas e para a qualidade do vínculo de trabalho
Especificação do teste*	<p>À luz das evidências recolhidas na avaliação, é um <i>straw-in-the wind test</i> (hipóteses futuras).</p> <p>Com base nas evidências previstas na tabela anterior, é um <i>straw-in-the wind test</i> (hipóteses futuras).</p>
Argumentação	<i>Straw-in-the wind test</i> (hipóteses futuras) – como nos mecanismos anteriores as evidências obtidas na análise de impacto contratual não são suficientes nem necessárias para confirmar a hipótese, ou seja, a sua obtenção nem valida o mecanismo, nem sugere a validação de eventuais hipóteses alternativas. Por outro lado, face às evidências previstas, este mecanismo apresenta alguns desafios a nível empírico pois é difícil a recolha de evidências que permitam comprovar inequivocamente que as medidas levam a uma redução dos riscos financeiros associados à mobilização de novos recursos humanos, a um aumento de confiança nas capacidades dos desempregados e assim à eliminação de barreiras à sua contratação. Encontrar evidências que existe esta causalidade torna o mecanismo mais plausível, mas podem existir inúmeros fatores explicativos que influenciam a confiança nas capacidades dos desempregados e que possibilitem a eliminação de barreiras à contratação.

Fonte: elaboração própria

Nota: Uma vez que este é um exemplo, optou-se por se efetuar a especificação do teste em dois momentos, uma vez que igualmente existem duas linhas de distinta de quando são aplicados os testes: Beach e Pedersen (2013) defendem que os testes devem ser aplicados antes de se recolherem as evidências de forma a orientar o processo de recolha, seleção de fontes e desenho de instrumentos; por sua vez no artigo do CDI (2015) é referido que estes devem ser aplicados após a recolha das evidências uma vez que existem pesquisas que não permitem ter uma ideia concreta de que evidências podem vir a ser obtidas.

Passo 5. Conclusão do processo

Face às evidências recolhidas, observa-se que não existe a confirmação nem a refutação da existência de nenhum dos mecanismos hipotéticos desenhados para a avaliação. Apesar das evidências indicarem que a participação nas tipologias de operação estágios e apoios à contratação levam a um

aumento da empregabilidade, a causa através da qual tal acontece não é analisada. Tal como já foi referido no capítulo anterior e é amplamente discutido na literatura, as análises contrafactuais (baseadas na análise da regularidade) encontram-se no “extremo oposto” à ACC no que se refere à análise da natureza da causalidade (Beach e Pedersen, 2013; CDI: 2015, Befani e Mayne, 2014). Por outro lado, os dados utilizados para a realização da Análise de Impacto Contrafactual, sendo bases de dados do IEFP e da Segurança Social não contemplaram variáveis que permitiriam aferir evidências relacionadas com os conceitos elencados nos mecanismos.

Quando se analisa face à evidência prevista os testes aos mecanismos 1, 2 e 4 são mais robustos, uma vez que se as evidências esperadas estiverem presentes, é confirmada a existência do mecanismo. No seguimento desta análise passar-se-ia para a construção dos instrumentos de recolha de informação e posterior recolha e análise das evidências, tal como referido nos passos anteriores. Após esta recolha, poder-se-ia repetir os testes para comprovar o poder confirmatório das evidências dos mecanismos.

No caso dos mecanismos 3 e 5 que apresentam o teste mais fraco, e uma vez que o objetivo é aumentar os níveis de certeza e singularidade, estes poderiam ser revistos no sentido de os tornar mais detalhados mediante uma maior concretização dos conceitos utilizados, bem como poderiam ser analisados à luz de outras teorias de forma a aumentar o poder inferencial.

6. Síntese Final

Esta etapa final do exercício de reporte procura situar alguns elementos conclusivos resultantes do percurso de estudo realizado.

Começamos por evidenciar que o campo da avaliação das políticas públicas em Portugal é, ainda, apesar de relevantes progressos verificados, deficitário de investimento na apropriação pela prática dos recursos metodológicos que o campo científico tem produzido. Esta perspetiva é explícita no que diz respeito à avaliação baseada na teoria que constitui campo de fundo do presente estudo.

Dão maior expressão à utilidade de um maior investimento metodológico o plano da configuração do exercício em sede de “encomenda” e o da capacidade de operacionalização, ou seja, da prática. São domínios interrelacionadas, mas com expectativa de progressos próprios que importa encorajar.

O recurso à Teoria da Mudança tem sido impulsionado no contexto da avaliação dos FEEI, como bem o demonstra o roteiro analítico precedente, e produziu um “choque metodológico” na prática da avaliação que tem alimentado o debate e alguma polémica. Tratando-se de uma certa rutura na prática de avaliação vigente era de esperar que assim fosse. Importará identificar as opções que podem permitir uma maior consolidação desta aposta.

A primeira refere-se ao processo de planeamento. A aplicação da Teoria da Mudança a estratégias de avaliação é significativamente beneficiada pela existência de referenciais de planeamento que mobilizem elementos-chave do seu referencial teórico. A eficácia de construir um referencial para a avaliação baseado na TdM que assuma um quadro de planeamento mais estruturado é evidente. O caminho feito nos FEEI neste domínio tem sido algum, mas, ainda, limitado. A exigência de um maior investimento em sede de planeamento das políticas é uma prioridade. A prática de planeamento que a própria Comissão Europeia promove no contexto, por exemplo, dos projetos de cooperação constitui um modelo que permite, simultaneamente, identificar a margem de progresso existente e cooptar soluções metodológicas.

Nesta ótica, importa construir as TdM na fase de planeamento da programação por forma a possibilitar a identificação atempada do racional da política estrutural bem como a sua coerência com as questões de avaliação.

Na ausência de um maior investimento em sede de planeamento, a opção por (re)construir as TdM como passo introdutório para a avaliação surge como incontornável. Contudo, a opção de o incluir como passo inicial dos processos de avaliação mostra-se uma solução difícil de operacionalizar e que tem produzido ganhos limitados. Esta dificuldade é ampliada pelo facto de, por circunstância (a dimensão dos programas operacionais) e uso (a prática consolidada ao longo do tempo) as intervenções objeto de avaliação (a encomenda) terem uma dimensão muito significativa e, como atestado na análise, uma complexidade elevada. Ora, uma e outra tornam a construção do referencial teórico de partida exercícios exigentes e morosos.

Assim, duas opções se colocam para uso cumulativo:

- i. O de colocar a construção das teorias de suporte à avaliação como uma etapa intermédia entre a programação e a avaliação, ou seja, como segunda etapa do processo (incompleto) de planeamento, permitindo que os exercícios de avaliação tenham como ponto de partida o

enunciado estratégico para validar. Esta opção compreende a vantagem de poder contribuir para a progressiva capacitação dos processos de planeamento por poder, e dever, ser suportada num amplo quadro de participação.

- ii. A de promover exercício de avaliação com uma configuração mais circunscrita que permita, em sede do próprio curso da avaliação, construir a abordagem teórica e a forma de metodologicamente lhe dar resposta. Esta opção reforça as condições para que a aplicação metodológica seja melhor sistematizada e mais detalhada.

No mesmo passo, a opção por antecipar e detalhar mais a definição dos roteiros teóricos permitirá analisar e discutir com maior detalhe a configuração dos exercícios de avaliação a desenvolver e como esta pode apoiar o robustecimento da prática, nomeadamente no que diz respeito ao plano metodológico. Esta evolução contribuirá para que a TdM em termos reais, concretos e práticos, com vista a implementar melhorias para o seu uso eficaz, identificando quais as oportunidades e limitações existentes.

Esta evolução é uma condição útil, mas não suficiente, à valorização dos exercícios de avaliação baseados na TdM. A análise produzida enfatiza, também, o desafio que a apropriação das abordagens metodológicas disponíveis representa para a comunidade de práticas. A reação a um processo que se revelou muito pesado e complexo sugere insuficiências no modo como este foi promovido, mas, igualmente, a dificuldade na mobilização e uso das estratégias metodológicas. A capacitação pela prática e pode ser complementada com abordagens mais estruturadas à aprendizagem metodológica, criando condições para uma utilização mais afinada dos métodos e a um debate mais informado sobre a prática. A estratégia de consulta à comunidade de práticas (stakeholders do processo) evidencia bem esta realidade.

A caracterização das intervenções acolhidas no âmbito dos FEEI explicita a sua natureza complexa e a exigência que coloca sobre o recurso à abordagem com base na TdM. São várias as implicações que se lhe associam.

A complexidade das intervenções reforça a utilidade de se apostar na TdM como estratégia para melhor compreender os referenciais programáticos subjacentes às políticas e garantir a sua avaliação. No entanto, não é linear a forma de o fazer. A resposta à complexidade remete para a valorização da análise da causalidade e esta para a atenção aos mecanismos. Contudo, como foi evidenciado pelo exercício de aplicação do *process tracing*, fazê-lo tende a exigir uma grande individualização da cadeia de programação e um significativo detalhe da análise. Isto será insuscetível de ser alcançado quando o campo da avaliação é amplo (macro).

O exercício de aplicação do *process tracing* mostrou que a recolha de evidências sobre o funcionamento de um determinado mecanismo poderá não ter elevada complexidade, mas beneficia com a especificação à posterior das opções metodológicas a considerar. Esta solução requer um objeto relativamente delimitado para que a avaliação seja manejável. A resposta às condições de funcionamento do mecanismo e a recolha de evidência associada permitirá, porém, um escrutínio bastante válido das políticas e uma consequente capacidade de incorporar valor no seu planeamento.

O exemplo estudado mostra como o recurso a abordagens contrafactuais é limitado como recurso explicativo. No caso dos instrumentos de política analisados, os mecanismos identificados (e não

estudados) podem influir de forma muito diversa e com intensidade diferentes os resultados alcançados. Como decorreu da própria avaliação desses instrumentos, a utilidade de os modelar em alinhamento com resultados de política no domínio da qualidade do emprego constitui uma prioridade. Para fazê-lo não bastará ter uma medida dos ganhos em termos de emprego (dado pelo exercício contractual). Fará falta saber como estes são gerados. O exercício incompleto de aplicação do *process tracing* mostra a sua utilidade para esse efeito e a importância de criar condições para aprofundar a análise da causalidade na avaliação das políticas públicas. A avaliação realista constitui-se como uma referência de grande utilidade neste plano de abordagem.

Contudo, outras estratégias devem ser consideradas na intensificação do uso da TdM no âmbito da avaliação dos FEEI. Uma ênfase excessiva de métodos detalhados de análise dos mecanismos conduz à inviabilização de análises mais integradas das políticas, o que deve ser evitado.

Assim, importará combinar diferentes planos de avaliação. Os micro com maior poder explicativo dos mecanismos e os macros com maior margem para explorar analiticamente os quadros lógicos da programação. Com efeito, a TdM é passível de incidir sobre uma ou mais das dimensões da programação consoante o critério definido pelos atores de avaliação. A mesmo tempo, importará ter em consideração que a TdM não requer a cobertura exhaustiva de todas as dimensões da programação. Importa definir os princípios que presidem a essa seleção e a sua flexibilidade face à geometria variável da programação.

A contribuição das teorias encaixadas (*nested theories*) afigura-se como recurso metodológico a contemplar. Estas consagram vários planos de análise e permitem combinar a elaboração e uso de várias teorias da mudança (de diferente alcance). No plano mais macro, a validação dos pressupostos da causalidade exige soluções menos especificadas, havendo que promover inferências mais indiretas legitimadas em quadros participados de consulta.

Em síntese, o foco nas dinâmicas de causalidade na avaliação dos instrumentos de política financiados pelos FEEI é uma condição requerida pelos atributos de complexidade que os caracteriza e que sugerem a necessidade de compreender a forma como atuam e o contexto em que o fazem. Contudo, a diversidade dos instrumentos objeto de avaliação e a utilidade de olhar para as dinâmicas de causalidade, combinando a capacidade de produzir leituras integradas da programação com a de pormenorizar o detalhe dos mecanismos inerentes às políticas avaliadas justifica a procura de soluções metodologicamente diferenciadas. Nesta ótica, o planeamento à medida das encomendas e processos de avaliação constitui um recurso útil à afirmação da TdM no âmbito da avaliação dos FEEI.

Referências Bibliográficas

Agência para o Desenvolvimento e Coesão (2016), Orientações para o acompanhamento da implementação das recomendações das avaliações do Portugal 2020 (follow-up), Julho.

Agência para o Desenvolvimento e Coesão (2019), Plano Global de Avaliação do Portugal 2020, dezembro. AD&C (2016), Orientações para o planeamento e preparação das avaliações do Portugal 2020, Dezembro.

Amândio, A. S. (2022) A avaliação dos fundos europeus em Portugal. *Public Sciences & Policies / Ciências e Políticas Públicas. Public Sciences & Policies / Ciências e Políticas Públicas*, 8(1).

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: mechanisms and theory building in evaluation. *American journal of evaluation*, 31(3), 363-381.

Barnes, Marian, Matka, Elizabeth, Sullivan, Helen (2003). "Evidence, Understanding and Complexity – Evaluation in Non-linear Systems". Sage, Vol 9, 265-284. London.

Barbrook-Jonhson, Castelline, Brian, Hills, Dione (2021). "Policy evaluation for a complex world: Practical methods and reflections from the UK Centre for the Evaluation of Complexity across the Nexus". Sage, Vol 27, 4-17.

Barbrook-Jonhson, Proctor, Amy, Giorgi, Sara, Phillipson, Jeremy (2020). "How do policy evaluators understand complexity". Sage, Vol 26, 315-332.

Barroso, Sergio (Coord) (2022) Roteiro para a capacitação do ecossistema dos fundos da política de coesão para o período 2021-2027, relatório final, CEDRU, EY, AD&C

Beach, D. e Pedersen, R.B. (2013), *Process Tracing Methods: Foundations and Guidelines*, Ann Harbor, University of Michigan Press.

Befani, B., & Mayne, J. (2014) "Process tracing and contribution analysis: A combined approach to generative causal inference for impact evaluation". *IDS Bulletin*, 45(6), 17-36.

Blamey, Avril, Mackenzie, Mhairi (2007). "Theories of Change and Realistic Evaluation". Sage. Vol 13, 439-455. London.

Brousselle, Astrid & Buregeya, Jean-Marie (2018) "Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation". *Evaluation* 24 (2), pp.153–168.

Cartwright, N., & Hardie, J. (2012) *Evidence-based policy: Doing it better. A practical guide to predicting if a policy will work for you*. Oxford, UK: Oxford University Press. <http://dx.doi.org/10.1093/acprof:osobl/9780199841608.001.0001>

CDI (2015), “Applying Process Tracing in Five Steps”, *Practice Paper Number 10 Annex*, Centre for Development Impact, April 2015.

CDI (2015a), “Straws-in-the-wind, Hoops and Smoking Guns: What can process tracing Offer To impact Evaluation”, *Practice Paper Number 10*, Centre for Development Impact, April 2015.

Chen H. T. (2006) A theory-driven evaluation perspective on mixed methods research. *Research in the Schools*, Vol. 13(1).

Chen, H. & P. Rossi (1989) “Issues in the Theory-driven Perspective”, *Evaluation and Program Planning* (12)4: 299–306.

Chen, H. T. (1980) The theory-driven perspective [Special issue]. *Evaluation and Program Planning*, 12, 299-306.

Chen, H. T., & Rossi, P. H. (1980) The multi-goal, theory-driven approach to evaluation: A model linking basic and applied social science. *Social Forces*, 59, 106-122.

Chen, H.T. (1990) *Theory-driven evaluations*. Newbury, CA: Sage.

Chris L. S. Coryn; Lindsay A. Noakes; Carl D. Westine; & Daniela C. Schroter (2011) A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009, *American Journal of Evaluation* 32(2) 199-226, DOI: 10.1177/1098214010389321

Coryn, Chris L. S.; Noakes, Lindsay A.; Westine , Carl D.; Schröter, Daniela C. (2011). “A Systematic Review of Theory Driven Evaluation Practice From 1990 to 2009”. *American Journal of Evaluation* 32 (2), pp. 199 226.

Costa, António Firmino da (1996) “Ciência e reflexividade social: relações entre ciência e sociedade segundo um inquérito aos investigadores portugueses”, em Maria Eduarda Gonçalves (coord.), *Ciência e Democracia*, Lisboa, Bertrand Editora, pp. 199-221.

Dionízio, S. & Rodrigues, D. (2022) A estratégia de avaliação do Portugal 2020. Uma perspetiva segundo a teoria da mudança. *Coleção Políticas e Territórios*, Caderno Temático, 3:36.

Directorate-General for Regional and Urban Policy (2021) COMMISSION STAFF WORKING DOCUMENT Performance, monitoring and evaluation of the European Regional Development Fund, the Cohesion Fund and the Just Transition Fund in 2021-2027 SWD(2021)198

Donaldson, S. I. (2007) *Program theory-driven evaluation science*. New York, NY: Lawrence Erlbaum.

Etzkowitz H, Leydesdorff L (1998) The endless transition: a “Triple Helix” of university–industry–government relations, introduction to a theme issue. *Minerva* 36:203–208

European Commission (2008, 2013a) EVALSED: The resource for the evaluation of Socio-Economic Development.

European Commission (2012) *Theory-based Evaluation*. Based on material produced for DG Regional Policy by Frans L. Leeuw.

European Commission (2013b) EVALSED sourcebook: Method and Techniques.

European Commission (2015) Guidance Document on Evaluation Plans – Terms of Reference for Impact Evaluations – Guidance on Quality Management of External Evaluations, fevereiro.

European Commission (2018) Employment, Social Affairs & Inclusion Programming period 2014-2020 Monitoring and Evaluation of European Cohesion Policy European Social Fund Guidance document.

European Commission (2018) Monitoring and Evaluation of European Cohesion Policy. Guidance Document, European Social Fund. Programming period 2014-2020, agosto.

European Commission (2019) Commission Staff Working Document. Synthesis of the Findings of the Evaluations of European Structural and Investment Funds Programmes. Strategic reports.

European Commission (2021) Performance, monitoring and evaluation of the European Regional Development Fund, the Cohesion Fund and the Just Transition Fund in 2021-2027, Comission Staff Working Document, julho.

Funnell, S., & Rogers, P. (2011) Purposeful program theory: Effective use of theories of change and logic models. San Francisco, CA: Jossey-Bass.

Gates, E. & Dyson, L. (2017) “Implications of the Changing Conversation About Causality for Evaluators”, *American Journal of Evaluation*, 38 (1), 29–46.

Glennan, S, (1996) “Mechanisms and the nature of causation”, *Erkenntnis*. Disponível em: https://digitalcommons.butler.edu/facsch_papers/27

Guba, E. G. & Lincoln, Yvonna S. (1989) Fourth Generation Evaluation, ISBN: 9780803932357, Sage Publications.

Hart, M. (2007) Evaluating EU regional policy. How might we understand the causal connections between interventions and outcomes more effectively? *Policy Studies*, Vol.28, n°4.

J. P. Connell, A. C. Kubisch, L. B. Schorr and C. H. Weiss (eds) (1995) *New Approaches to Evaluating Community Initiatives: Volume 1, Concepts, Methods, and Contexts*. Washington, DC: The Aspen Institute.

James, C. (2011) Theory of change review (Report commissioned by Comic Relief). Retrieved from <http://mande.co.uk/blog/wp-content/uploads/2012/03/2012-ComicRelief-Theory-of-Change-Review-FINAL.pdf> Johnson, N., Guedenet, H., & Saltzman, A. (2014).

Jesuíno J C (1993) Comunidade científica: dinâmicas e fronteiras, *Actas do II Congresso Português de Sociologia*, II, 48-57.

Jesuíno, J C (coord.); Amancio, L; Avila, P; Carapinheiro, G.; Costa, A F; Machado, F L; Patrício, M T; Stoleroff, A; Vala, (1995) *J. A comunidade científica portuguesa nos finais do século XX*, Celta.

Kuhn, T. S. (1962) *The Structure of Scientific Revolution*, Chicago, The University of Chicago Press

Leeuw F. (2003) Reconstructing Program Theories: Methods Available and Problems to be Solved, in «American Journal of Evaluation», n. 24(1), pp. 5-20.

Leeuw F., Vaessen J. (2009) Impact Evaluations and Development: NONIE guidance on impact evaluation, Network of Networks on Impact Evaluation (NONIE).

Leeuw Frans L. (2012) Theory Based Evaluation, http://ec.europa.eu/regional_policy/information/evaluations/pdf/impact/theory_impact_guidance.pdf

Mackie, J. L. (1974). The cement of the universe: A study of causation. Oxford, UK: Oxford University Press.

Mamede, R. P., e Pereira, H. (2018). The impact of EU-funded direct subsidies on several dimensions of firm performance in Portugal: 2008- 2015. Working paper: DINÂMIA'CET

Mayne J. (2001) Addressing Attribution through Contribution Analysis: Using Performance Measures Sensibly, in «Canadian Journal of Program Evaluation», n. 16(1), pp. 1-24.

Mayne, J. (2012). Contribution analysis: Coming of age? Evaluation, 18 (3), 270 – 280. <http://dx.doi.org/10.1177/1356389012451663>

Mayne, J. (2015) Useful Theory of Change Models, Canadian Journal of Program Evaluation 30.2, 119–142 doi: 10.3138/cjpe.230

Mayne, J., & Johnson, N. (2015) Using theories of change in the cgair research program on agriculture for nutriDon and health. *Evaluation*, 21(4), 407–428 Pawson, R. (2007). Causality for beginners. Não publicado, disponível em: [hPps://eprints.ncrm.ac.uk/id/eprint/245/](http://eprints.ncrm.ac.uk/id/eprint/245/)

Merton, R. (1973/1942) The Sociology of Science. Chicago, The University of Chicago Press

Merton, R. K. (1968/1949). *Social theory and social structure*. The Free Press.

Patton, M. Q. (2008) Utilization-focused evaluation (4th ed.). Thousand Oaks, CA: Sage.

Pawson, R. (2002) Evidence-based Policy: The Promise of 'Realist Synthesis', in «Evaluation», n. 8(3), pp. 340-358.

Pawson, R., Tilley, N., & Tilley, N. (1997) "In with the new: introducing the Scientific Realism", The new rules of realistic evaluation", Realistic evaluation. Londres. Sage

PO ISE (2022) Avaliação intercalar do PO ISE Anexos do relatório final Volume I. Disponível em: <https://poise.portugal2020.pt/avaliacao-e-resultados>

PO ISE (2022a) Avaliação intercalar do PO ISE. Relatório Final. Disponível em: <https://poise.portugal2020.pt/avaliacao-e-resultados>

Polany, M. (1962) The Republic of Science, its political and economic theory, Chicago: Roosevelt University.

Riché, M. (2013) Theory Based Evaluation: A wealth of approaches and an untapped potential. *Capturing effects of projects and programmes. Lund: Studentlitteratur*, 85.

Rogers, P. J. (2008) Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14, 29-48.

Stame, N. (2004) Theory-based evaluation and types of complexity. *Evaluation*, 10, 58-76

Stein, D., & Valters, C. (2012) Understanding ‘theory of change’ in international development: A review of existing knowledge. Asian Institute and the Justice and Security Research Programme. http://www.theoryofchange.org/wp-content/uploads/toco_library/pdf/UNDERSTANDINGTHEORYOFChangeSteinValtersPN.Pdf

Steel, Daniel (2003). “Social Mechanisms and Causal Inference” Sage, Philosophy of the Social Sciences. Vol. 34, 55-78.

Stern, E. (2020) Editorial, *Evaluation* Vol. 26(4) 401–403 DOI:10.177/135638902096442

Stern, E., Stame, N.; Mayne, J.; Forss, K.; Davies, R.; Befani, B. (2012) “Broadening the range of designs and methods for impact evaluations”. *Department for International Development Working Paper 8*.

Vogel, I. (2012) Review of the use of ‘theory of change’ in international development. Department for International Development. Retrieved from http://www.oxfamblogs.org/fp2p/wp-content/uploads/DFID-ToC-Review_VogelV4.pdf

Weiss, C. (1972) *Evaluation*. Englewood Cliffs, NJ: Prentice Hall.

Weiss, C. (1995) “Nothing as Practical as Good Theory: Exploring Theory-based Evaluation for Comprehensive Community Initiatives for Children and Families”, in J. P. Connell, A. C. Kubisch, L. B. Schorr and C. H. Weiss (eds) *New Approaches to Evaluating Community Initiatives: Volume 1, Concepts, Methods, and Contexts*. Washington, DC: The Aspen Institute.

Weiss, C. (1997) “Theory-based evaluation: Past, present and future”, *New directions for evaluation: a publication of the American Evaluation Association*, Vol.76(114), p.41-55 ISSN: 1097-6736, 1534-875X.

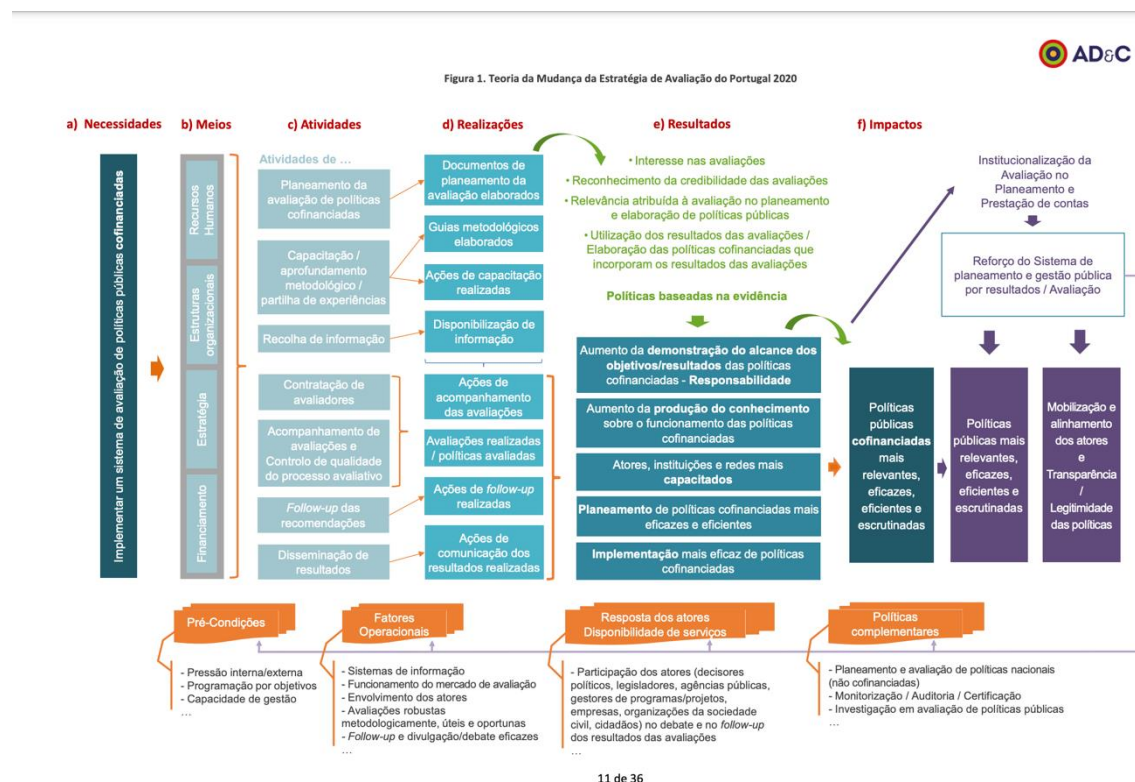
Weiss, C. (1998) “Understanding the programme” and “Developing measures”, *Evaluation (2nd Editions)*, New Jersey, Prentice Hall: 46-70; 114-150.

Weiss, C. (2000) “Which Links in Which Theories Shall We Evaluate?”, in P. J. Rogers, T. Hacsí, A. Petrosino and T. A. Huebner (eds) *Program Theory in Evaluation: Challenges and Opportunities, New Directions for Evaluation*, 87. San Francisco, CA: Jossey-Bass.

Anexo

A recomendação da simplificação da TdM foi feita com base nos esquemas inframencionados.

Figura 11 – Teoria da Mudança da estratégia de avaliação do Portugal 2020



11 de 36

Fonte: Dionizio e Rodrigues (2022)

Figura 12 – Tipos de avaliações e critérios de avaliação



Fonte: PGA PT2020. Adaptado de CE (2014), *Guidance Document on Monitoring and Evaluation – CF and ERDF – Concepts and Recommendations*