

Mónica Catarina do Adro Lopes

**AVALIAÇÃO NAS ORGANIZAÇÕES DO TERCEIRO SETOR:
CONTEXTOS, DISCURSOS E PRÁTICAS**

Volume I

Tese de Doutoramento em Sociologia - Cidades e Culturas Urbanas, orientada pela Professora Doutora Virgínia Ferreira e apresentada à Faculdade de Economia da Universidade de Coimbra.

Outubro de 2017



UNIVERSIDADE DE COIMBRA



FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Mónica Catarina do Adro Lopes

AVALIAÇÃO NAS ORGANIZAÇÕES DO TERCEIRO SETOR: CONTEXTOS, DISCURSOS E PRÁTICAS

Volume I

Tese de Doutoramento em Sociologia - Cidades e Culturas Urbanas, apresentada à
Faculdade de Economia da Universidade de Coimbra para obtenção do grau de Doutora.

Orientadora: Professora Doutora Virgínia Ferreira

Coimbra, 2017

*Para a minha filha, Maria,
que nasceu com esta tese.*

AGRADECIMENTOS

Desenvolver uma tese de doutoramento é sempre um processo que envolve sentimentos ambivalentes de aproximação e distanciamento, de descoberta e ocultação, de entusiasmo e descrença, de rumo e de deriva. Mas é, sobretudo, um percurso através do qual nos (desc)construímos e consolidamos a nossa identidade. Percorrer este caminho não teria sido possível sem as disponibilidades e as amizades de todas as pessoas e entidades que ao longo destes anos me apoiaram. Nomeio algumas, mas a minha gratidão está longe de se esgotar nestas referências.

As minhas primeiras palavras de agradecimento dirigem-se às pessoas e organizações que por razões diversas tornaram possível esta investigação. Às pessoas que, em representação de organizações de base e de cúpula terceiro setor, de programas comunitários, e especialistas e profissionais na área da avaliação, gentilmente se disponibilizaram para colaborar neste trabalho – os seus contributos constituem o essencial desta investigação, e por isso esta tese também é deles/as. À Fundação para a Ciência e Tecnologia pelo financiamento concedido. À Faculdade de Economia da Universidade de Coimbra e ao Centro de Estudos Sociais por terem aceite, enquadrado e acolhido este doutoramento. Um agradecimento especial ao CES, que me proporcionou as melhores condições logísticas e académicas para prosseguir o estudo.

À minha orientadora, Virgínia Ferreira, mestre e companheira em muitos projetos, pela segurança, pelo apoio nos momentos de maior ansiedade, pelas palavras sensatas e amigas, pelo incentivo e confiança que sempre depositou no meu trabalho. Foi com ela que dei os primeiros passos nesta aprendizagem, num trajeto que se sedimenta mas não culmina neste trabalho.

À Sílvia Ferreira, pelos muitos caminhos que me mostrou e portas que me abriu em momentos de conversa, pela imensa generosidade, pela inspiração.

Aos/às colegas e amigos/as do CES, em cuja companhia tudo se tornou tão mais leve. Cúmplices das ansiedades, das inseguranças, dos desencantos, mas também dos devaneios, das fugas e dos muitos risos. Um agradecimento especial à Filipa, ao Fernando, à Hélia, à Rafaela, ao Daniel e à Denise, em cujas palavras e gestos encontrei energia, conforto e confiança para seguir em frente em momentos de desassossego.

Ao Gato, mão invisível que me segurou com firmeza (aqui chegámos!). À Cláudia e ao Vasco, que há já longos anos me sustentam em períodos de gáudio e apuro. À Catarina, em cujo abraço encontrei alegria maior nesta fase final. Ao meu irmão, Bruno, cujos passos acompanharam e impulsionaram sempre os meus.

Este trabalho foi financiado pela bolsa de doutoramento SFRH/BD/74823/20110 da Fundação para a Ciência e Tecnologia, cofinanciada por fundos nacionais, através do Ministério da Educação e Ciência, no âmbito do programa POPH, participado pelo Fundo Social Europeu (FSE).



RESUMO

A avaliação tem assumido uma importância crescente nas organizações do terceiro setor (OTS) em resultado dos desafios originados por uma nova relação com o Estado e outras entidades financiadoras, e das transformações ocorridas no seio do próprio setor. Articulado um conjunto amplo de perspetivas teóricas e analíticas, radicadas em campos do conhecimento diversos, o estudo propôs-se perscrutar motivações, práticas e usos da avaliação nas OTS portuguesas, desvelando lógicas e racionalidades associadas às modalidades de avaliação adotadas, e refletindo sobre as suas implicações na identidade do setor. A estratégia de pesquisa assentou na análise interpretativa das fontes documentais e dos registos orais das entrevistas semiestruturadas realizadas a 45 atores dos campos da avaliação e do terceiro setor, e na aplicação de um inquérito por questionário *online* a representantes de 575 OTS. O estudo é, grosso modo, revelador da ainda frágil cultura de avaliação nas organizações estudadas, em que a subscrição efusiva da retórica da avaliação não ecoa no investimento que é feito na integração de práticas consistentes e na criação de estruturas próprias de avaliação potenciadoras de processos de aprendizagem ao serviço das suas missões. A análise permitiu desvelar a implicação e relevância de uma diversidade de fatores, quer associados ao ambiente interno, quer associados ao contexto ambiental e institucional, na composição deste retrato geral, assim como na definição de tipologias de avaliação e diferentes perfis organizacionais. Ao mesmo tempo que evidencia a centralidade das pressões institucionais na configuração das práticas e representações sobre avaliação, especialmente as associadas ao financiamento, o estudo sublinha a relevância do papel dos fatores contingenciais internos e da cultura organizacional na mediação dos efeitos do ambiente externo na internalização da avaliação.

Palavras-chave: terceiro setor; avaliação; medição; desempenho; impacto; prestação de contas; aprendizagem organizacional.

ABSTRACT

Recent changes both within third sector organizations (TSO) and in the relationship between TSO and the State and other funding bodies have increased the importance of evaluation practices. Making use of different theoretical and analytical perspectives, this research aims to study the motivations, the practices and the uses of evaluation by Portuguese OTS. This is intended to expose the logics and different rationalities associated with the evaluation practices adopted and to disclose its implications upon the identity of the TSOs. The research strategy has been based on secondary data analysis, 45 semi-structured interviews with different agents of the evaluation and the third sector fields, and an on-line questionnaire to 575 OTS representatives. The analysis of collected data revealed that the actual evaluation culture is still very fragile in the studied organizations. Despite the effusive discourses towards the adoption of an evaluation culture within TSO, there isn't a consistent adoption of evaluation practices or the development of infrastructures to support them, which would allow organizational learning and mission achievement. The study has shown the influence of a diversity of internal and external factors upon the reported evaluation practices, as well as upon the definition of evaluation typologies and organization's profiles. The analysis has highlighted the significance of institutional pressures in the configuration of both evaluation practices and the ways they are envisaged, especially those associated with funding. However, this study has also underscored the relevance of internal contingency factors and TSO organizational culture in mediating the effects of the external environment towards on the internalization of evaluation.

Keywords: third sector; evaluation; measurement; performance; impact; accountability; organizational learning.

LISTA DE SIGLAS E ABREVIATURAS

- ADL – Associação de Desenvolvimento Local
- ANIMAR – Associação Portuguesa para Desenvolvimento Local
- BSC – *Balanced Scordcard*
- CASES – Cooperativa António Sérgio para a Economia Social
- CE – Comissão Europeia
- CERCI – Cooperativa p/Educação e Reabilitação de Crianças Inadaptadas
- CLDS – Contrato Local de Desenvolvimento Social
- CNES – Conselho Nacional para a Economia Social
- CNIS – Confederação Nacional das Instituições de Solidariedade
- CONFECOOP – Confederação Cooperativa Portuguesa
- CPCCDRD – Confederação Portuguesa das Coletividades de Cultura, Recreio e Desporto
- CPF – Centro Português de Fundações
- CSSES – Conta Satélite da Economia Social
- DAC - Development Assistance Committee (Comité de Apoio ao Desenvolvimento)
- DGERT – Direção Geral do Emprego e das Relações de Trabalho
- EAPN – *European Anti Poverty Network* (Rede Europeia Anti Pobreza)
- EQUASS – *European Quality in Social Services*
- EU – União Europeia
- EUA – Estado Unidos da América
- FEDER – Fundo Europeu de Desenvolvimento Regional
- FEEI – Fundos Europeus Estruturais e de Investimento
- FENACERCI – Federação Nacional de Cooperativas de Solidariedade Social
- FSE – Fundo Social Europeu
- INSCOOP – Instituto António Sérgio para o Setor Cooperativo
- IPSS – Instituição Particular de Solidariedade Social
- ISO – *International Organization for Standardization* (Organização Internacional para Padronização)
- ISS – Instituto de Segurança Social
- ML – Modelo(s) Lógico(s)
- MSC – *Most Significant Changes* (Mudanças Mais Significativas)
- OCDE – Organização para a Cooperação e Desenvolvimento Económico
- ONG – Organizações Não Governamental
- ONGA – Organização Não Governamental para o Desenvolvimento
- ONGD – Organização Não Governamental para o Desenvolvimento
- ONGDM – Organização Não Governamental para os Diretos das Mulheres
- ONGPD – Organizações Não Governamental das Pessoas com Deficiência

OTS – Organização do Terceiro Setor
PIC EQUAL – Programa de Intervenção Comunitária EQUAL
PO – Programa Operacional
POEFDS – Programa Operacional O Emprego, Formação e Desenvolvimento Social
POPH – Programa Operacional Emprego e Formação
QCA – Quadro Comunitário de Apoio
QREN – Quadro de Referência Estratégico Nacional
QUAR – Quadro de Avaliação e Responsabilização
SIADAP – Sistema Integrado de Gestão e Avaliação do Desempenho na Administração Pública
SROI – *Social Return on Investment* (Retorno Social do Investimento)
SWOT – *Strengths, Weaknesses, Opportunities and Threats* (Forças, Fraquezas, Oportunidades e Ameaças)
TS – Terceiro Setor
UMP – União das Misericórdias Portuguesas
UMP – União das Mutualidades Portuguesas
USAID – United States Agency for International Development (Agência dos Estados Unidos para o Desenvolvimento)

Avaliação nas Organizações do Terceiro Setor: Contextos, discursos e práticas

ÍNDICE

INTRODUÇÃO	1
PARTE I: ASPETOS CONCETUAIS E ESTRATÉGIAS ANALÍTICAS	9
Capítulo 1. Fundamentos e perspetivas no estudo da avaliação.....	11
1.1. Definições, objetivos e funções.....	11
1.2. Paradigmas, perspetivas e teorias de avaliação: enquadramento da diversidade de abordagens	21
1.2.1. Paradigma <i>experimental</i> : abordagens clássicas	25
1.2.2. Paradigma <i>construtivista</i> : abordagens naturalistas	27
1.2.3. Paradigma <i>transformativo</i> : abordagens emancipatórias.....	29
1.2.4. Paradigma <i>pragmático</i> : abordagens utilitaristas	31
1.3. Abordagens managerialistas à avaliação	34
1.3.1. Avaliação <i>versus</i> controlo e auditoria.....	36
1.3.2. Entre diferentes tradições disciplinares: Avaliação <i>versus</i> medição do desempenho	45
Capítulo 2. Avaliação nas organizações do terceiro setor: contextos e abordagens ..	53
2.1. Governação, desestatização e projetificação das políticas sociais	53
2.2. O panorama atual do terceiro setor em Portugal	63
2.3. Desafios da gestão das organizações do terceiro setor: dilemas da profissionalização	73
2.4. Abordagens à eficácia organizacional e desempenho das organizações do terceiro setor: contributos da teoria das organizações	81
2.4.1. A avaliação do desempenho e do impacto do terceiro setor enquanto processo socialmente construído	89
2.5. Relações de accountability e avaliação nas organizações do terceiro setor	92
2.5.1. Perspetivas teóricas sobre a adoção de sistemas de <i>accountability</i>	96
2.5.2. Quadros interpretativos da adesão e uso da avaliação: entre influências externas e internas	101
2.6. Cultura de avaliação e aprendizagem organizacional.....	109

Capítulo 3. Estratégia metodológica: um plano misto de pesquisa 115

3.1. Objetivos.....	115
3.2. Dimensões analíticas	116
3.3. Hipóteses de trabalho	124
3.4. As fontes e estratégias de recolha e análise da informação: um desenho metodológico misto.....	126
3.4.1. Entrevistas.....	128
3.4.2. Inquérito por questionário às organizações do terceiro setor	133
3.4.3. Análise documental.....	137

PARTE II: O ESTUDO EMPÍRICO 141

Capítulo 4. Discursos e práticas de avaliação no terceiro setor português 143

4.1. Tipos de discurso sobre avaliação: científico, associativo e gestor.....	147
4.2. A subscrição efusiva da retórica da avaliação	155
4.3. Os diferentes enquadramentos da avaliação nas organizações do terceiro setor: da gestão estratégica à gestão de projetos.....	159
4.3.1. Os projetos como enquadramento privilegiado dos exercícios formais de avaliação.....	159
4.3.2. A prevalência da medição do desempenho na avaliação à escala organizacional	165
4.3.3. A centralidade da conformidade com normas e procedimentos nos processos de monitoria externa	172
4.3.4. Compreendendo a implementação de práticas de gestão, controlo e avaliação: perfis de organizações do terceiro setor	175
4.4. Decompondo as práticas avaliativas: métodos, conteúdos e intervenientes...183	
4.4.1. Metodologias de avaliação: entre a informalidade das abordagens qualitativas e apreciativas e a conformidade com os requisitos de descrição e quantificação.....	183
4.4.2. Âmbitos de incidência e conteúdos da avaliação: o enfoque nos processos e realizações e o desafio da orientação para os impactos	200
4.4.3. Participantes na avaliação: o fraco envolvimento dos stakeholders....	218
4.5. Padrões de utilização dos resultados das avaliações	227
4.5.1. Combinando métodos e usos da avaliação: implicações diferenciadas na cultura avaliativa	230

4.5.2. A influência do envolvimento dos <i>stakeholders</i> nos usos da avaliação	235
4.5.3. Compreendendo os usos da avaliação: entre a institucionalização e a estratégia	239
4.6. Síntese: a prevalência das lógicas managerialistas e tecnocráticas de medição do desempenho.....	244
Capítulo 5. Capacidade <i>para fazer</i> avaliação: contextos e competências.....	251
5.1. Estruturas organizacionais de suporte à avaliação	253
5.1.1. Recursos humanos	253
5.1.2. Recursos financeiros	260
5.2. Ambiente interno: receptividade e suporte da liderança	263
5.3. O ambiente externo: iniciativas de suporte e capacitação	269
5.4. Capacidade para fazer avaliação: retratos organizacionais	284
5.5. Síntese: a debilidade dos contextos, recursos e estruturas de avaliação.....	291
PARTE III: CONCLUSÕES.....	295
Capítulo 6. Perfis e modelos de avaliação nas organizações do terceiro setor	297
6.1. Perfis de avaliação nas organizações do terceiro setor	297
Perfil I – Avaliação enquanto ferramenta de gestão estratégica	304
Perfil II – Avaliação enquanto ferramenta de projeção externa	308
Perfil III – Avaliação como fardo ou dispersão	312
Perfil IV – Avaliação como mecanismo de responsabilização interna e promoção local	316
Capítulo 6.2. Práticas heterogêneas em ambientes diversos: caminhos para o desenvolvimento de uma cultura avaliativa	318
CONSIDERAÇÕES FINAIS	329
REFERÊNCIAS BIBLIOGRÁFICAS	339

ÍNDICE DE FIGURAS, GRÁFICOS, QUADROS E TABELAS

Figuras

Figura 1. Funções da avaliação	18
Figura 2. Tipos de uso da avaliação	20
Figura 3. Relações entre fatores explicativos da avaliação, uso da avaliação e eficácia organizacional.....	107
Figura 4. Enquadramento do processo de avaliação.....	144
Figura 5. Cadeia de valor da intervenção das organizações.....	201
Figura 6. Dimensões analíticas da capacidade <i>para fazer</i> avaliação.....	253
Figura 7. Dimensões e subdimensões associadas à cultura avaliativa nas OTS	298

Gráficos

Gráfico 1. Tipos de discurso sobre avaliação: científico, associativo e gestor.....	149
Gráfico 2. Palavras-chave mais comuns e “interlinguagem”	153
Gráfico 3. Perceções sobre impacto do POPH e EQUAL na organização*	163
Gráfico 4. Tipo de informação de monitorização/avaliação recolhida	213
Gráfico 5. Comunicação dos resultados de avaliação	220
Gráfico 6. Níveis de participação de <i>stakeholders</i> na avaliação.....	221
Gráfico 7. Uso dos resultados da avaliação (N=556)	228
Gráfico 8. Constrangimentos à avaliação	259
Gráfico 9. Principal fonte de financiamento da avaliação (%).....	261
Gráfico 10. Competências em avaliação segundo pertença a plataformas/redes internacionais	277
Gráfico 11. Fontes de informação sobre avaliação	278
Gráfico 12. Perceções sobre suporte recebido para implementação da avaliação (POPH <i>versus</i> EQUAL)*	280

Quadros

Quadro 1. O contínuo da avaliação forma/informal	16
Quadro 2. Características das principais correntes/tradições de avaliação.....	23
Quadro 3. Comparação entre instrumentos de gestão/avaliação	42
Quadro 4. Comparação entre avaliação e medição de desempenho	49
Quadro 5. As metodologias da intervenção social pública	58
Quadro 6. Pressupostos e críticas das abordagens à eficácia/desempenho organizacional	85
Quadro 7. Características dos tipos de prestação de contas	96

Quadro 8. Contributos teóricos para a compreensão das motivações e usos da avaliação	103
Quadro 9. Práticas de gestão, controlo e avaliação: perfis de OTS.....	177
Quadro 10. Características dos paradigmas quantitativo e qualitativo	183
Quadro 11. Balanço anual de atividades de OTS	193
Quadro 12. Suportes da recolha de informação: perfis de OTS	196
Quadro 13. Indicadores de realização (output) e resultado do POPH	204
Quadro 14. Usos da avaliação: Perfis de OTS	240
Quadro 15. Dimensões da capacidade de avaliação das organizações.....	252
Quadro 16. Obstáculos à avaliação: perfis de OTS	286
Quadro 17. Competências internas e externas em avaliação: perfis de OTS.....	289
Quadro 18. Perfis de avaliação nas OTS	301
Quadro 19. Visão geral das características dos tipos de avaliação	325

Tabelas

Tabela 1. Variáveis de caracterização das OTS entrevistadas	131
Tabela 2. Tipologia das organizações da amostra, por comparação com a Conta Satélite da Economia Social (CSES) (%)	135
Tabela 3. Variáveis de caracterização das OTS inquiridas	135
Tabela 4. Práticas de gestão, controlo e avaliação*	158
Tabela 5. Frequência do envolvimento em atividades de gestão e avaliação segundo tipo de informação de avaliação recolhida com maior frequência	190
Tabela 6. Correlações entre práticas de controlo, gestão e avaliação e níveis de participação global, interna e externa na avaliação	223
Tabela 7. Frequência das atividades de controlo, gestão e avaliação segundo os tipos de usos da avaliação	231
Tabela 8. Correlações entre tipo de informação de avaliação privilegiada e perceções sobre avaliação	234
Tabela 9. Correlações entre tipos de usos da avaliação e perceções sobre avaliação.....	235
Tabela 10. Correlações entre usos da avaliação e níveis de participação de <i>stakeholders</i> na avaliação.....	238
Tabela 11. Obstáculos à avaliação (Componentes extraídas da ACCP)	285

INTRODUÇÃO

As organizações do terceiro setor têm assumido uma relevância crescente um pouco por todo o mundo. Os múltiplos papéis que têm desempenhado, quer de reforço da coesão social e económica, através do combate às diversas formas de exclusão social e de marginalidade económica e da criação de riqueza, quer de promoção da cidadania ativa assente na participação cívica e política, têm contribuído para o seu crescimento exponencial desde a década de 1980. Em parte, a afirmação do terceiro setor como ator político e económico foi alimentada tanto pelo desenvolvimento da responsabilidade do Estado pelo bem-estar dos cidadãos como pela retração do papel do Estado na provisão direta de bem-estar e, ainda, pelas novas lógicas da governação. Surgem novos processos, mecanismos e atores cujas relações estão constantemente a mudar, assim como as suas fronteiras de atuação, admitindo-se que a governação está a tornar-se cada vez mais difusa (Ferreira, 2009, 2012). As funções crescentes na provisão de bens e serviços e na criação de emprego, assim como o reconhecimento por parte do Estado (e de instâncias europeias e internacionais) das capacidades e potencialidades do terceiro setor na promoção da participação dos cidadãos na construção de uma sociedade mais justa e democrática e no alcance de níveis adequados de bem-estar, posicionam estas organizações como interlocutores e parceiros privilegiados das políticas públicas.

Este posicionamento, possibilitando às organizações do terceiro setor (OTS) ocupar um lugar proeminente nos processos de governação societal, sujeita-as também a uma maior visibilidade e exposição ao escrutínio público por parte de diversos *stakeholders*. A parceria com o Estado na prossecução do interesse geral, designadamente através contratualização para o fornecimento de serviços e da participação em programas/projetos sociais, carrega novas exigências de produção e de prestação de contas, nomeadamente mediante o estabelecimento de requisitos de monitorização e avaliação das atividades. Num contexto marcado por novos modelos de intervenção social emergentes do Estado-Providência ativo, baseados, nomeadamente, na abordagem por projeto e na partilha de responsabilidades pela realização dos objetivos com outros parceiros

(Hespanha, 2008), a avaliação surge como elemento indispensável na contratualização entre Estado e terceiro setor tendo em vista a implementação de políticas sociais.

A orientação para a nova gestão pública no Estado, prescrevendo uma cultura de resultados centrada no cumprimento de objetivos, intensificou este controle, estendendo assim à relação entre o Estado e as OTS novas exigências de qualidade na prestação dos serviços e requisitos de maior eficácia e eficiência na intervenção social (Ferreira, 2012). O desenvolvimento de políticas regionais promovidas pela União Europeia (EU), ao colocar em marcha o mecanismo dos Fundos Estruturais (FEDER e FSE), tem constituído um motor de desenvolvimento no que diz respeito às reformas da administração pública, em geral, e de avaliação das políticas públicas, em particular, ao obrigar os estados-membros a realizar atividades de avaliação sistemáticas que justifiquem os programas de apoio. Estas exigências acabam por ser transpostas para a relação do Estado com as OTS, por intermédio das relações que estabelecem para a implementação daqueles programas. Os fundos estruturais europeus constituem, de resto, uma importante fonte de financiamento para muitas instituições e para algumas respostas na área social, por via da implementação de projetos.

Desde o início do movimento mais recente no sentido de tornar os organismos públicos mais eficientes e adequados (Osborne e Gaebler, 1992) que têm vindo a aumentar as pressões sobre as organizações do terceiro setor no sentido de também elas se tornarem mais eficazes. Esta tendência geral, que tem sido chamada de "movimento de prestação de contas¹", refere-se à noção de que as organizações do terceiro setor, e as pessoas que as conduzem, devem ser mais "responsáveis" perante os públicos que é suposto servirem e perante quem as financia (Murray, 2005: 347). É esta pressão do ambiente em que se move o setor – financiadores, utentes, reguladores e outros intervenientes – que, segundo muitos/as autores/as (Kearns, 1996; Murray, 2005; Gibelman e Gelman, 2008; Ebrahim, 2010), tem sido a principal causa do crescente interesse das organizações na avaliação. Por outro lado, tendo aumentado em número e visibilidade em muitas partes do mundo ao longo das últimas décadas, as organizações têm também sido atingidas por inúmeros

¹ Neste trabalho os termos "prestação de contas", "*accountability*" e "responsabilização" são usados indistintamente.

escândalos, amplamente divulgados, relacionados com a falta de prestação de contas ou o respetivo enviesamento/viciação. Tanto dirigentes como pessoal técnico têm sido objeto de acusações de irregularidades que vão desde a má gestão de recursos e utilização de fundos para fins pessoais a abusos sexuais e fraude (Gibelman e Gelman, 2001, 2004; Cohen, 2007)². Têm contribuído para a erosão da confiança pública nas OTS (Gibelman e Gelman, 2004, 2001) e para o abandono da visão destas organizações enquanto "pílulas mágicas" preocupações com situações relacionadas com remunerações indevidas da direção, elevados custos de gestão, operacionalização e captação de recursos, acumulação de riqueza, mercantilização e fracasso no apoio a quem precisa, (Najam, 1996: 340). Tais preocupações têm levado pessoas ligadas à intervenção e investigação no campo social a refletir sobre a acomodação de mecanismos de prestação de contas, nomeadamente os baseados no desempenho das organizações, como é o caso da avaliação.

Paralelamente, no seio do próprio terceiro setor tem ocorrido um amplo movimento no sentido da "mercantilização" ou "comercialização" das suas atividades e da "profissionalização" da sua gestão (Hvenmark, 2013; Maier, Meyer e Steinbereithner, 2014). A maior orientação para o mercado, a par da maior presença de profissionais na estrutura de gestão, introduz dinâmicas no sentido de tornar o setor mais eficiente e racional, mediante a adoção de uma série de sistemas de gestão, medição e avaliação do desempenho, que correspondem frequentemente a modelos e práticas importados do setor empresarial (Hwang e Powell, 2009). Estas forças combinam-se com a crescente popularidade do empreendedorismo social (Dees, 1998; Dees e Anderson, 2006) e da filantropia de risco/de impacto (*venture philanthropy*) (John, 2006; Moody, 2008), sendo que ambos se inspiram fortemente em métricas e práticas do setor empresarial. Enquanto a criação de valor social e a mensuração/avaliação dos resultados e impacto social criado se encontram entre os pilares estruturantes do empreendedorismo social, nas novas lógicas de financiamento do setor associadas ao investimento de impacto, a medição do impacto social assume-se como critério essencial para a tomada de decisão de quem investe.

² Em Portugal, destacam-se as polémicas em que esteve envolvida a Liga Portuguesa Contra o Cancro (*Público*, Maio 1996), Abraço (*Público*, Março 2000 e Julho 2005) e, mais recentemente, o Centro Português Para a Cooperação (CPPC) (*Observador*, Setembro de 2014), a Obra Diocesana de Promoção Social (*DN*, Abril de 2015), a Casa do Gaiato (*Observador*, Março 2016) e os Centros Paroquias de Carvalhais e de Centro Paroquial de S. Martinho das Moitas, em São Pedro do Sul (*TSF*, Fevereiro de 2016).

Estas forças têm contribuído para a crescente relevância da avaliação e da medição do desempenho/do impacto nas OTS na literatura académica internacional, motivando, por um lado, a proliferação de obras conceituais e teóricas e de estudos empíricos sobre a condução de processos de medição/avaliação nas organizações e, por outro lado, a explosão de metodologias e ferramentas para a avaliação do desempenho/impacto nas OTS.

Em Portugal, só muito recentemente se começou a discutir a necessidade de prestação das contas e de demonstração de resultados e impactos nas OTS. Daí que o trabalho teórico e empírico sobre a temática da avaliação nas OTS seja ainda escasso no nosso país.

O presente estudo pretende contribuir para colmatar esta lacuna, propondo-se conhecer motivações, práticas e usos da avaliação nas OTS a operar em Portugal, desvelando lógicas e racionalidades associadas às modalidades de avaliação adotadas. Procura-se carrear contributos para a resposta a quatro questões essenciais associadas à compreensão da avaliação nas OTS: o quê? (que práticas de avaliação são conduzidas nas organizações); como? (com que recursos – técnicos, financeiros e materiais); porquê? (que contextos, fatores e motivações influenciam a sua adoção); e para quê? (com que propósitos e utilidade). Num plano mais específico, pretende-se compreender em que medida as práticas de avaliação se têm ou não vindo a constituir instrumentos de governação e gestão estratégica das organizações, assim como delinear o papel que as entidades financiadoras, em particular as associadas a programas cofinanciados por fundos estruturais europeus, têm tido nesse processo.

Tais desígnios são desafiados pela amplitude e complexidade do objeto de estudo, cuja análise convoca uma grande diversidade de quadros analíticos oriundos de áreas do conhecimento distintas, com tradições, linguagens, perspetivas e abordagens diferenciadas. Todavia, se a diversidade de olhares disciplinares sobre a avaliação nas OTS constitui sem dúvida um grande contributo, ela representa também um enorme desafio teórico e analítico, dada a dispersão do conhecimento que tem resultado dos diferentes *backgrounds* de produção. Por exemplo, os entendimentos e concetualizações sobre “avaliação” na literatura sobre teoria de avaliação (que em geral se associa a programas) e

na literatura na área da gestão ou da teoria organizacional são claramente distintos. A avaliação não constitui uma prática unificada, possui diversas raízes, que fazem emergir uma pluralidade de fontes e de filosofia da avaliação, nomeadamente: as iniciativas nos anos 60 da *Great Society* nos Estados Unidos; a inovação educativa e dos *curricula* nas escolas; a gestão por objetivos; a gestão com base nos resultados; a investigação participativa no desenvolvimento rural e comunitário, entre outras.

Fazer comunicar e integrar no estudo perspectivas radicadas em tradições disciplinares distintas constitui uma missão desafiante mas necessária para uma leitura abrangente do fenómeno nos contextos observados. Com efeito, neste trabalho recorreu-se a um conjunto alargado de contribuições teóricas e analíticas radicadas em diversas áreas científicas – sociologia, economia e ciência política –, e diversos campos de estudo – teoria de avaliação, teoria das organizações, gestão estratégica, estudos sobre terceiro setor, estudos sobre eficácia, etc.).

Tais contributos, colhidos através de uma exaustiva revisão da literatura nas diferentes áreas do conhecimento, alimentaram a construção do quadro de referência teórico e analítico que se apresenta na primeira parte do trabalho, composta por três capítulos. No primeiro capítulo apresenta-se um esboço geral dos paradigmas e das perspectivas que têm marcado o campo da avaliação, quer em termos da sua delimitação concetual quer em termos da definição de modelos (prescritivos ou descritivos) das atividades de avaliação. Suportada sobretudo na literatura sobre teoria de avaliação, delinheio alguns dos contributos para a definição do conceito, por referência à sua função, aos seus métodos e aos seus objetivos, e sistematizo, numa perspectiva histórica, as teorias de avaliação com base nos seus pressupostos ontológicos, epistemológicos e metodológicos. Articulam-se ainda as abordagens influenciadas por desenvolvimentos teóricos e empíricos sobre avaliação originários da gestão empresarial com particular afinidade com os novos modelos de gestão pública.

O segundo capítulo situa a avaliação no contexto das organizações do terceiro setor. Começo por fundamentar a relevância da avaliação nas novas tendências de governação, nas novas metodologias de intervenção social e nas atuais pressões para a racionalização e profissionalização do terceiro setor, e traço brevemente o panorama atual

do setor em Portugal. Prossegue-se, numa linha mais teórica, com a discussão sobre eficácia organizacional, percorrendo os principais modelos para explorar o desempenho e o impacto das OTS e questionando o estabelecimento de um modelo de avaliação comum para o setor, a partir da visão da eficácia e do impacto enquanto processos socialmente construídos. No ponto seguinte enquadro a avaliação no “movimento de prestação de contas”, discutindo as principais perspetivas teóricas sobre a adoção de sistemas de *accountability*, a partir das quais delinheiro quadros interpretativos da adesão e uso da avaliação. Por fim, adotando uma perspetiva integrada sobre prestação de contas e aprendizagem organizacional, exploro as relações entre avaliação e aprendizagem organizacional, focando o papel da cultura organizacional na construção da capacidade avaliativa.

No último capítulo da primeira parte do estudo - o terceiro - são referenciadas as principais opções metodológicas do estudo. Especifica-se a natureza da investigação e apresentam-se os principais elementos da pesquisa, desde os objetivos que a nortearam, passando pelo mapa conceptual definido para capturar a evidência empírica, as hipóteses esboçadas, aos elementos metodológicos e trabalhos realizados no sentido de operacionalizar a investigação.

Os quarto e quinto capítulos enformam a segunda parte do trabalho, dedicada à análise e reflexão sobre a informação produzida e sistematizada. O quarto capítulo, mais longo, centra-se sobretudo na resposta às questões “o quê”, “porquê” e “para quê”. Por um lado, são exploradas as modalidades de monitorização e avaliação adotadas pelas organizações do terceiro setor abrangidas no estudo (em termos da sua natureza, enquadramentos, propósitos, métodos, intervenientes e conteúdos), partindo sobretudo do olhar sobre o seu desenho metodológico e sobre as lógicas inerentes às estratégias prosseguidas (científica, associativa ou gestonária). Por outro lado, são analisados os propósitos e motivações que presidem à adoção de práticas de avaliação, e desvelados padrões de utilização dos seus resultados (e do processo em si), equacionando a influência do ambiente organizacional e de fatores de contingência associados à identidade e estrutura organizacionais, e perscrutando dinâmicas de prestação de contas e de aprendizagem organizacional.

O quinto capítulo – refletindo sobre a questão “como” – centra-se na capacidade das OTS *para fazer* avaliação, a partir da análise das dimensões associadas aos recursos disponíveis para a realizar, humanos e organizacionais mais amplos, e das dimensões associadas ao contexto organizacional interno (cultura organizacional) e externo (suporte técnico e financeiro de atores relevantes no campo do terceiro setor).

A terceira parte do estudo, conclusiva, é composta por um capítulo, no qual é feita uma análise global das práticas e representações sobre avaliação nas OTS, mediante a definição de perfis e modelos de avaliação, compostos a partir de um conjunto de dimensões e sub-dimensões que influenciam a cultura avaliativa das organizações e que compõem o modelo analítico de referência mobilizado ao longo da investigação.

O trabalho termina com algumas considerações finais em torno das principais conclusões possibilitadas pelo estudo, das limitações que afetaram o seu desenvolvimento e das pistas de investigação deixadas em aberto.

PARTE I: ASPETOS CONCETUAIS E ESTRATÉGIAS ANALÍTICAS

Capítulo 1. Fundamentos e perspectivas no estudo da avaliação

1.1. Definições, objetivos e funções

São múltiplas as significações associadas ao conceito de “avaliação” e o emprego que dele se faz é muito diverso, ilustrando bem a complexidade que lhe está associada. Cada uma das definições incide num ou noutro aspeto segundo o paradigma em que se enquadra, as perspectivas às quais se liga e à funcionalidade que se lhe atribui. Cada definição é, em certa medida, uma declaração de princípios na qual prevalecem certos valores que se pretendem concretizar.

Trata-se, pois, de um termo dúctil, com usos distintos e aplicado em vários contextos e realidades, desde o mundo do trabalho (avaliação de desempenho), à educação (avaliação de aprendizagens), à saúde (avaliação de serviços), à ação social (avaliação de programas/projetos). Os objetos das avaliações podem ser, pois, diversos: leis, produtos, serviços, organizações, pessoas, políticas, processos e projetos sociais, etc. A maioria das definições situa-se no âmbito da avaliação de programas, mais do que de políticas ou de organizações. É, aliás, frequente as perspectivas e definições, assim como as teorias e os modelos, de avaliação que se anunciam amplas, conterem elementos relacionados especificamente com avaliação de programas. No entender de Michael Scriven (2013: 172), esta circunstância traduz uma **falácia “geocêntrica”** de assumir que a avaliação é avaliação de programas³. Esta falácia não será alheia ao contexto de forte disseminação da avaliação após a Segunda Guerra Mundial (sobretudo nos EUA), muito associado à multiplicação de programas (sociais, de desenvolvimento urbano, de habitação social, de combate ao insucesso escolar, etc.) e à necessidade de determinar os efeitos das experiências da *Great Society* dos anos 60 (nomeadamente dos programas de combate à pobreza). Neste trabalho adoto a abordagem que Scriven designa **“heliocêntrica”**, que toma como ponto central a *avaliação* e como subaéreas a avaliação de programas, de projetos, de pessoal, de organizações, de serviços, de produtos, etc.

³ Encontramos nesta definição um exemplo desta perspectiva: Avaliação de um conjunto estruturado de intervenções (atividades ou projetos) para atingir determinados objetivos específicos a diversas escalas. (*Adaptado de DAC, 2010*).

Aparte o viés “geocêntrico” de muitas definições de avaliação, um domínio essencial em que elas diferem é o das componentes que incluem (Shaw, Greene e Mark, 2006). Algumas definições centram-se na *função* genérica que a avaliação serve. A definição funcional mais comum indicará que a avaliação envolve julgamentos de valor, indicações de mérito, valia e relevância de algo.

Outras definições elegem como aspeto central a especificação dos respetivos *objetivos* (por exemplo, fornecer informação para o desenho de políticas ou para a melhoria dos programas, serviços, ou organizações). Dada a multiplicidade de objetivos da avaliação existem variações, dependendo do fim que se pretende dar às avaliações (desde informar as decisões a fomentar a democracia). Além disso, algumas definições especificam com maior centralidade o *método* de avaliação. Elencam-se de seguida algumas propostas que encontramos na literatura.

Capucha *et al.* definem avaliação de uma forma abrangente, sobretudo por referência aos seus objetivos:

Formas deliberadas e racionalizadas de colocação de questões e problemas relativos aos processos de decisão e execução de programas, políticas, projetos e investimentos, com recurso a sistemas organizados de reflexão crítica a partir de informações recolhidas no decurso do acompanhamento de programas, políticas, projetos e investimentos. O objetivo último será permitir que as pessoas e as instituições envolvidas na conceção, planeamento, gestão e execução julguem o seu trabalho e os resultados obtidos e aprendam com eles. (Capucha *et al.*, 1996: 11)

Cohen e Franco (Cohen e Franco, 2006: 73) propõem uma definição com base na sua função genérica, sugerindo que “avaliar é fixar o valor de uma coisa; para ser feito requer um procedimento mediante o qual se compara aquilo a ser avaliado com um critério ou padrão determinado”.

Na sua definição funcional, Isabel Guerra releva, por seu turno, a avaliação enquanto medida:

Avaliar é sempre comparar com um modelo – medir – e implica uma finalidade operativa que visa corrigir ou melhorar. O padrão ou modelo a partir do qual se avalia é, em última instância, um valor de referência que, numa situação de planeamento, se encontra geralmente fixado, a partir do diagnóstico da situação inicial, nos objetivos e metas definidos.” (Guerra, 2002: 185)

A autora ressalva que a avaliação não é uma simples medição dos resultados finais, e que os dados recolhidos e organizados monograficamente não constituem por si só uma

avaliação, mas é necessário interpretá-los a partir de um padrão, que é definido não só em termos científicos mas também em termos político-ideológicos (Guerra, 2002: 186).

No seu trabalho sobre avaliação de políticas públicas, Cuesta e Martínez (2010: 13) começam por avançar uma definição genérica de avaliação, sustentando que “avaliar é apreciar, analisar, valorar e julgar os resultados e impactos de uma intervenção, a fim de melhorar a sua qualidade”. Segundo as autoras, a avaliação constitui um instrumento integrado de análise das políticas públicas, de carácter multidisciplinar e de recente especialização profissional, que tem como objetivo apreciar a eficácia dos programas ou políticas públicas, comparando os seus objetivos e impactos previstos e os meios disponíveis, emitindo um juízo de valor. Mais do que um instrumento de conhecimento *per se* é um processo guiado por uma série de exigências e valores: rigor, imparcialidade, transparência, atenção aos diferentes pontos de vista e prevalência do interesse geral (*ibidem*: 13-14). Apesar de emitir um juízo de valor, a avaliação não tem a recriminação ou a sanção como fim último, mas sim uma maior compreensão da ação pública.

À semelhança de Cuesta e Martínez, também Daniel Stufflebeam e Chris Coryn (2014: 14) – proeminentes teóricos no domínio da avaliação, em particular na avaliação de programas – estendem a definição genérica de avaliação no sentido de abranger um conjunto de valores e critérios que consideram “particularmente importantes numa sociedade livre e democrática”: “A avaliação é o processo sistemático de delimitação, obtenção, reporte e aplicação de informação descritiva e de juízo acerca do mérito, valor, probidade, viabilidade, segurança, significância e/ou equidade de um objeto.” Nesta definição operacional, além dos valores, os autores acrescentam à componente funcional os principais tipos de informação e de tarefas abrangidas pela avaliação de programas.

No campo específico da investigação social aplicada e da implementação de programas/projetos, as definições avançadas para o termo revestem-se de um carácter eminentemente prático, diretamente associado à apreciação dos resultados e impactos de um programa relativamente ao objetivo de resolução de determinados problemas. Assim, Kosecoff e Fink definem avaliação como “um conjunto de procedimentos para julgar os méritos de um programa e fornecer uma informação sobre os seus fins, as suas expectativas, os seus resultados previstos e imprevistos, o seu impacto e os seus custos” (1982: 15).

Também a definição avançada por Donna Mertens (1998: 219), pioneira da abordagem transformativa em avaliação, remete para o domínio dos programas: "a avaliação é a investigação sistemática do mérito ou valor de um objeto (programa) com a finalidade de reduzir a incerteza na tomada de decisão."

No âmbito da avaliação de programas sociais, outra definição de referência é a avançada por Peter Rossi (2013: 109-110), que define a pesquisa de avaliação como a "aplicação sistemática de métodos de pesquisa em ciências sociais para a apreciação da concepção, desenho, implementação e utilidade dos programas de intervenção social".

A avaliação é concebida, na linha destas definições, como parte da ciência social aplicada. Nesse sentido, faz uso de um grande espectro de teorias sociais, conceitos e métodos de investigação, sendo as mesmas regras que governam a recolha de informação válida e fidedigna, fundamental para a ciência (Stockmann, 2011: 18).

Na sua revisão sobre o estado da arte das conceitualizações de avaliação, Michael Scriven (2013: 167-179) contesta as concepções da avaliação enquanto ciência social aplicada e até enquanto ciência. No entender do autor, embora "haja uma boa dose de ciência na avaliação, ela não é somente uma ciência, assim como não é apenas uma arte ou apenas um ramo da literatura" (*ibidem*: 168). Tal como a lógica, a matemática, o *design*, etc., a avaliação é, para Scriven, uma transdisciplina: uma disciplina de pleno direito que serve não apenas as ciências mas outras disciplinas, não se encontrando (como se encontra a ciência) vinculada ao estudo de um subconjunto de fenómenos. Ela serve um conjunto alargado de disciplinas não científicas, desde o direito, a história, a literatura, a tecnologia e a educação, até à dança, à ginástica e ao poker. Por isso não pode ser chamada ciência; pode ser, no máximo, considerada um elemento da ciência, como a lógica ou a matemática. Nesta sequência, o autor problematiza a designação dada à avaliação de ciência social aplicada, uma vez que ela é também (ao contrário da ciência social) "uma parte crucial da física, da química, da medicina e da matemática" e, tal como as outras transdisciplinas (e contrariamente às ciências sociais aplicadas), é uma disciplina autónoma (*ibidem*: 168).

A discussão semântica sobre avaliação conduzida por Scriven remete para a distinção entre **avaliação formal e avaliação informal**, delineada por diversos/as autores/as (Tassie, Murray e Cutt, 1998; Carman e Fredericks, 2008; Stufflebeam e Coryn,

2014), e que considero relevante para clarificar o entendimento sobre avaliação que adoto neste trabalho. Tassie, Murray e Cut (1998: 63), por exemplo, situam esta distinção na discussão metodológica sobre avaliação, diferenciando métodos formais de métodos informais. Enquanto com os primeiros se pretende que sejam "objetivos" e "racionais" e impliquem o estabelecimento explícito de critérios de avaliação formal, a designação de pessoas responsáveis pela aplicação desses critérios e a produção e análise de dados que lhes estão associados, os métodos não-formais remetem para conversas informais, impressões e reputação, constituindo bases para a avaliação não oficializada.

Na verdade, enquanto atividade informal, a avaliação existe em todos os lugares, sob muitas formas. A produção de juízos e a tomada de decisões a partir do que se observa é constante nos diversos domínios da vida. Em essência, avalia-se constantemente, quer se trate de tomar decisões sobre o trivial ou sobre o crítico (Posavac e Carey, 2003). Dependendo da natureza do objeto de avaliação, podem buscar-se alternativas, ler-se rótulos, consultar-se pessoas com experiência pertinente, formar-se grupos de trabalho para deliberar sobre questões avaliativas relevantes, consultar-se fontes de informação diversas, ou experimentar-se algo antes de o adquirir. Trata-se de iniciativas de avaliação apropriadas que se encaixam numa conceção de avaliação informal. Todavia, como notam Stufflebeam e Coryn (2014: 26), a condução de avaliações informais é propensa à recolha casual de informação, a dar crédito ou a usar propaganda e outras formas de desinformação, a erros de julgamento, a agir com base em velhas preferências ou preconceitos, a recorrer a informação desatualizada, a depender do *feedback* de uma amostra inadequada ou tendenciosa de clientes, ou a fazer escolhas onerosas. Em muitos casos, os procedimentos da avaliação informal não são sistemáticos, são pouco rigorosos e baseados em perspetivas enviesadas (Stufflebeam e Coryn, 2014: 26). As avaliações formais, por seu turno, têm um carácter sistemático e rigoroso: deverão ser relevantes, desenhadas e implementadas no sentido de controlar enviesamentos, documentadas, manter-se consistentes com padrões profissionais adequados, reportadas a quem de direito, e de outra forma úteis e defensáveis (*ibidem*: 26-27).

Quadro 1. O contínuo da avaliação forma/informal

Informal	←—————→	Formal
Casual		Deliberada
Espontânea		Requer tempo
Critérios são implícitos		Critérios são explícitos e claros
Obtenção rápida e pouco apurada de evidências		Evidências são sistematicamente recolhidas
O juízo é feito de forma “imponderada”		O juízo é feito de forma meticulosa

Fonte: Elaboração da autora.

De facto, como nota Scriven, a avaliação não é apenas uma ciência e, tal como a tecnologia ou outros fenómenos, não obedece necessariamente aos métodos das ciências. Mas quando falamos de avaliação formal, falamos da avaliação como (trans)disciplina que, tal como reconhece Michael Scriven (2013: 172), para justificar tal designação, recorrerá a abordagens e métodos que são válidos, dados bem estabelecidos e teorias testáveis ou de outro modo avaliáveis, ainda que, como faz notar o autor, se possa questionar se o sistema de controlo de qualidade do sistema de avaliação responde aos padrões mínimos das disciplinas clássicas (*ibidem*: 172-175), ou problematizar a robustez e validade (ou mesmo a existência) de normas e padrões profissionais no campo da avaliação.

Neste sentido, entendo neste trabalho a avaliação enquanto processo de pesquisa de reflexão estratégica, que vai além das reflexões espontâneas e intuitivas, exigindo a definição de objetivos, a identificação exaustiva dos elementos de informação a recolher, assim como o estabelecimento de um método de reflexão.

Em todo o caso, de acordo com o Stockmann (2011), o facto de ser voltada para a apreciação e para a aplicação prática constitui um pré-requisito essencial da utilidade da avaliação. Por isso, a avaliação não pode ser reduzida a aspirações puramente científicas, não tendo em regra como objetivo primeiro a produção de conhecimento teórico geral, mas sim a aplicação de procedimentos e conhecimentos científicos para dar suporte a questões relacionadas com clientes e grupos-alvo. Neste sentido, a pesquisa em avaliação distinguir-se-á por uma “dualidade” especial, já que é por um lado parte da ciência social empírica, utilizando as respetivas teorias e métodos, e por outro lado parte do processo político que é ele mesmo influenciado pelos seus resultados.

Para lá da investigação básica, o recurso da avaliação à metodologia científica destinar-se-á a contribuir para a resolução prática de problemas sociopolíticos ou organizacionais, empenhando-se propositada e sistematicamente no sentido de disponibilizar informação para os processos de tomada de decisão. Assim,

A pesquisa em avaliação desempenha uma espécie de equilíbrio entre o que é científico e a aplicação política e a aplicação prática. Por um lado, a avaliação é parte da ciência social empírica, e possui regras e padrões a que deve obedecer. Por outro lado, é orientada para a produção de resultados úteis para a melhoria da prática social (Stockmann: 19).

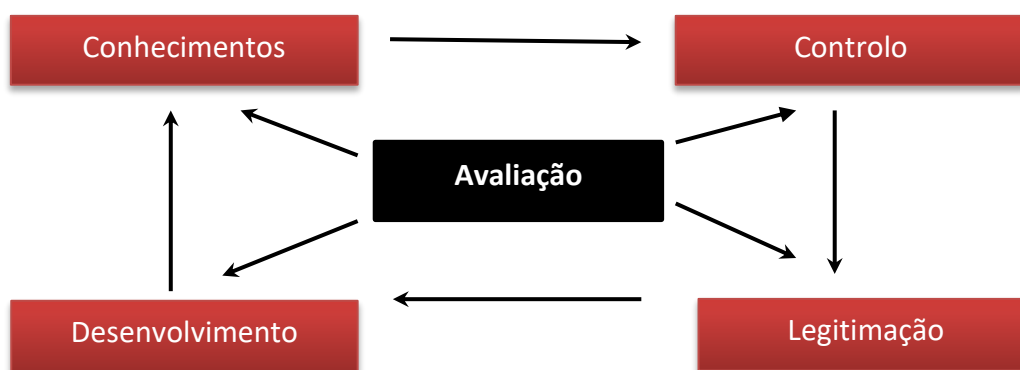
Apesar da amplitude e abrangência de algumas definições propostas, as diversas faces que a avaliação ostenta e a sua natureza dinâmica justificam que nenhuma definição tenha reunido o consenso. Além disso, os resultados e impactos da avaliação poderão ser muito distintos, dependendo do tipo de processo de avaliação seguido. Não obstante, Capucha *et al.* (1996: 11) sistematizam um conjunto de **contributos** mais frequentes:

- Melhoria do autoconhecimento das entidades, da forma como funcionam, se afastam e aproximam dos seus objetivos, como utilizam os recursos, que oportunidades de desenvolvimento estratégico possuem;
- Melhoria da condução e gestão de serviços, programas e projetos, com afinação dos mecanismos de autocorreção;
- Produção de informações sistematizadas que constituem memórias das instituições e dos dispositivos de intervenção e meio de reprodução dos projetos inovadores;
- Melhoria do processo de tomada de decisão, de definição de prioridades e de gestão estratégica;
- Identificação de variáveis mais facilmente alteráveis e das de mais difícil manipulação;
- Melhoria dos mecanismos de difusão de informação, de promoção da participação e de negociação nos processos de mudança;
- Formação de atores no domínio do julgamento de ideias, atividades e resultados e no domínio dos desempenhos profissionais (capacidades técnicas, sociais e relacionais);
- Prestação de contas e apuramento de resultados das atividades, desempenhos e projetos;

- Teste e afinação de teorias, conceitos, metodologias e conhecimentos sobre determinados assuntos, através da respetiva devolução aos agentes.

Estritamente associados aos contributos da avaliação estão os **objetivos** que presidem aos processos de avaliação. De facto, os resultados da avaliação podem servir vários propósitos que, embora possam ser separados analiticamente, estão intimamente ligados. Stockmann (2011: 26) distingue, *grosso modo*, quatro objetivos da avaliação (vide figura 1): (i) procurar conhecimento; (ii) exercer controlo; (iii) criar transparência e possibilidades de diálogo no sentido de produzir desenvolvimento, e (iv) legitimar ações/medidas implementadas.

Figura 1. Funções da avaliação



Fonte: Stockmann, Reinhard (2011) *A practitioner handbook on evaluation*. Cheltenham: Edward Elgar, Figura 2.2.

Especificamente:

(i) A avaliação deverá fornecer **conhecimentos** no sentido de informar e racionalizar as decisões. Poderá ser importante obter informação que permita aferir se determinada iniciativa está ou não a funcionar; quais as necessidades dos públicos-alvo e se estas estão ou não a ser satisfeitas, o grau de aceitação que a intervenção está a merecer; se as organizações promotoras estão ou não em condições de intervir com efetividade e eficiência; de que forma o contexto da intervenção se altera e de que modo isso tem impacto na intervenção; se os objetivos e impactos esperados são atingidos; que contributos a intervenção dá para a resolução dos problemas identificados, etc. A

informação é recolhida no sentido de criar conhecimento que será apreciado mediante os critérios de avaliação acordados ou especificados para a intervenção/iniciativa.

(ii) Sem a produção de conhecimento sobre o desenvolvimento de estruturas e processos, nenhuma avaliação poderá criar benefício. Todavia, nem sempre é o processo de decisão que encabeça a aplicação dos conhecimentos adquiridos, mas sim o **controle** (veja-se, por exemplo, Snibbe, 2006; Hoole e Patterson, 2008). Trata-se sobretudo de prestar contas, em que os critérios de sucesso como a efetividade, a eficácia, a eficiência, a aceitabilidade e sustentabilidade são aplicados com esse fim. Mesmo quando as avaliações não têm como fim primeiro o controle, toda a avaliação tem inerente, direta ou indiretamente, uma certa forma de controle, já que tem como regra perceber se as pessoas e entidades envolvidas numa intervenção desempenham ou não as tarefas que lhes estão atribuídas e cumprem as obrigações a que estão vinculadas.

(iii) Tanto as avaliações orientadas para o conhecimento como as orientadas para o controle produzem informações que podem ser utilizadas no **desenvolvimento** de determinada intervenção. Quando os resultados são divulgados, tem lugar um diálogo entre os diferentes tipos de *stakeholders* (entidades financiadora e promotora, grupos-alvo e outros). Nesta função da avaliação, predomina a *avaliação de processos*, a qual será utilizada para o desenvolvimento de outras intervenções. Esta função desempenha um papel central nas avaliações formativas.

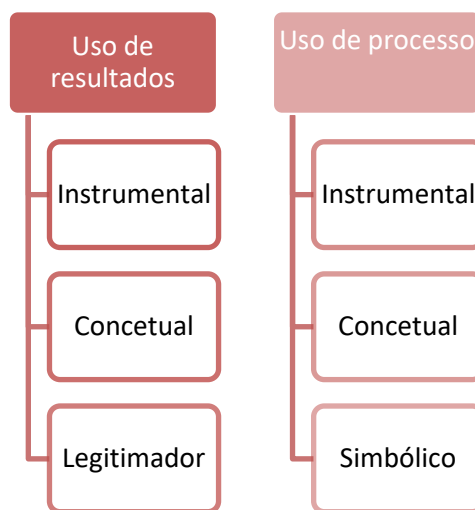
(iv) Uma outra função da avaliação é **legitimar** determinada intervenção ou medida. O conjunto de dados obtidos com a ajuda de uma avaliação oferece a possibilidade de fornecer evidências sobre os *outputs* e resultados obtidos, e de provar sobre a eficiência com que os fundos têm sido geridos as com que intervenções têm sido conduzidas. Num contexto em que os fundos são escassos, esta função da avaliação apresenta-se crescentemente relevante, uma vez que as organizações e respetivas intervenções competem frequentemente entre si e os *policy-makers* são obrigados a estabelecer prioridades e a fazer escolhas.

Também é frequente atribuir *funções táticas* às avaliações, nos casos em que se pretende usar os enunciados da avaliação para legitimar determinadas posições políticas quando, por exemplo, se pretende que determinada intervenção seja continuada ou, pelo

contrário, suspensa. Por outro lado, tornou-se “moda” utilizar a avaliação, enquanto símbolo de política moderna, sem que haja uma intenção séria de utilizar os seus resultados. Todavia, como salienta Stockmann (2011: 28), dificilmente este tipo de função tática pode ser compatibilizado com os reais propósitos da avaliação, podendo ser representada como o seu lado patológico.

As avaliações podem, na linha desta tipologia, ter diferentes funções, as quais não são independentes entre si, antes, estão intimamente ligadas. Com efeito, sem que a função de produção de conhecimento seja assegurada, nenhuma das outras funções poderá ser prosseguida. Todavia, também é certo que avaliações nas quais outras funções são prioritárias também produzem conhecimentos. O comprometimento com uma função primária preside à abordagem e determina a forma como a avaliação é desenhada e conduzida (*ibidem*: 28). Os objetivos da avaliação estão também intimamente ligados à utilização que se faz dos seus resultados (ou dos seus processos).

Figura 2. Tipos de uso da avaliação



Fonte: Alkin, Marvin C.; Taut, Sandy (2003) “Unbundling Evaluation Use” *Studies in Educational Evaluation*. 29 (3), 1–12, Figura 3.

A discussão sobre os objetivos e funções da avaliação alinha-se, pois, com o debate teórico em torno do **uso/utilização da avaliação**. Na revisão desta discussão conduzida por Alkin e Taut (2003) e retomada por autores/as como Fleischer e Christie (2009: 159) e Johnson *et al.* (2009: 378), são identificados dois domínios distintos de utilização: **uso de processo** e **uso de resultados**. Este último tem sido tradicionalmente dividido em três tipos: instrumental, concetual e simbólico. O uso instrumental acontece quando quem toma

decisões usa os resultados da avaliação para modificar o objeto da avaliação (neste caso, a avaliação é usada para produzir efeito na ação direta); o uso conceitual acontece quando os resultados da avaliação não induzem nenhuma ação direta, mas suscitam alterações em entendimentos conceituais específicos, decorrentes de mudanças na forma como quem utiliza a avaliação pensa sobre aspectos específicos da intervenção; e o uso simbólico remete para situações em que a simples existência da avaliação, e não qualquer aspecto dos seus resultados, é usada para persuadir ou convencer. Neste caso, a avaliação é usada tanto para justificar uma decisão que já foi tomada quanto para demonstrar que uma intervenção ou organização está aberta à avaliação, e por isso reforça a reputação de quem decide. O uso de processo é um conceito mais recente, definido por Patton (2012: 143) como "mudanças cognitivas, atitudinais e comportamentais em indivíduos, resultantes, direta ou indiretamente, do envolvimento no processo de avaliação". Ocorre quando as pessoas envolvidas na avaliação aprendem a partir do processo de avaliação em si ou fazem alterações nas intervenções, baseando-o no processo de avaliação e não apenas nos seus resultados. O uso da avaliação em função do seu processo pode também ocorrer conceitualmente, instrumentalmente ou simbolicamente (Alkin e Taut, 2003), ainda que estes tipos de utilização sejam comumente associados ao uso dos resultados da avaliação.

1.2. Paradigmas, perspectivas e teorias de avaliação: enquadramento da diversidade de abordagens

Considerando que os objetivos que presidem à avaliação podem variar grandemente e que tem havido mudanças importantes nos perfis e requisitos de avaliação ao longo da história, é compreensível a existência de uma grande variedade de abordagens e modelos, contruídos a partir de diferentes critérios, fundamentos e tradições. Não sendo possível dar conta do largo espectro de perspectivas e teorias sobre avaliação, tentarei agora sistematizar, numa perspectiva histórica, as várias abordagens existentes com base na sua orientação em termos metodológicos e de teoria da ciência.

A avaliação não tem sido alheia aos paradigmas científicos da investigação social já que durante muitos anos a sua legitimidade se baseou em critérios técnico-científicos

importados da investigação social. Apesar de posteriormente se ir desenvolvendo como (trans)disciplina com especificidades próprias e ir criando os seus próprios critérios, isso não impediu que continuasse imersa e no centro dos grandes debates científicos relativos à investigação social.

Por entre êxitos e fracassos, a avaliação consolida-se entre os anos 60 e 70, falando-se neste período de uma “indústria em crescimento”, estreitamente ligada ao positivismo e aos desenhos experimentais. O processo de avaliação, na sua primeira geração (que Guba e Lincoln (1989) apelidam de “mensuração” e Vedung (2010) de “vaga científica”), consistiu sobretudo na recolha de dados mensuráveis e factos quantificáveis, por forma a permitir comparações. No fundo, procurou-se adotar os métodos utilizados nas ciências naturais, usando-os para a pesquisa social empírica. Mas é ainda nos anos 60, no auge do positivismo experimentalista, que começa a emergir um novo paradigma que submete a forte revisão crítica o paradigma clássico-quantitativo dominante. Esta nova orientação científica baseada nos métodos qualitativos, com uma orientação mais pluralista e participativa, ver-se-á reforçada pelos processos de modernização dos Estados, pela nova consciência de cidadania e de aprofundamento democrático, e pelo próprio fracasso do experimentalismo americano para explicar por si só a complexidade da realidade (Guerra, 2002). É aqui que entra em cena aquela que Guba e Lincoln (1989) apelidam de “avaliação de quarta geração” e que Vedung (2010) designa de “vaga orientada para o diálogo”.

No quadro que se segue sintetizam-se as características das principais correntes da avaliação, em diferentes dimensões.

Quadro 2. Características das principais correntes/tradições de avaliação

	Positivista	Pragmática	Construtivista	Transformativa
Lógica	Científica - experimental	Utilitarista - Tomada de decisão e melhoria	Naturalista - Processos sociais	Emancipatória – Justiça social
Pressupostos teóricos	Existe uma única realidade que pode ser objetivamente documentada eliminando explicações alternativas Objetividade e neutralidade de quem avalia; observação distante e desapaixonada	Existe uma única realidade e cada indivíduo tem dela a sua interpretação própria As relações estabelecidas na avaliação são determinadas pelo que quem avalia considera ser apropriado a cada análise	Realidades múltiplas, socialmente construídas Ligação interativa entre quem dinamiza e participa na avaliação; valores são explicitados; diálogo e reflexão significativos para criação de conhecimento	Múltiplas realidades formatadas por valores sociais, políticos, culturais, económicos, étnicos, de género, deficiência e outros valores Conhecimento é socialmente e historicamente situado; a interação entre quem avalia interage e os <i>stakeholders</i> tem em conta quem se encontra em situação de desvantagem
Principais defensores	R. Tyler; D. Campbell, H. Chen; P. Rossi.; R. Boruch; T. Cook; S. Donaldson; L. Cronbach; W. Sadish	J. Wholey; D. Stufflebeam; M. Patton; M. Alkin; H. Perskill; J.B. Cousins; M. Q. Patton	R. Stake; Y. Lincoln; E. Guba; E. Eisner; I. Shaw; M. Parlett; D. Hamilton;	D. Mertens; J. Greene; B. MacDonald; B. Whitemore
Principais modelos	Avaliação por Objetivos; Avaliação Experimental ou Quase-Experimental; Avaliação Baseada em Teoria	Avaliação Focada na Utilização; Modelo CIPP; Avaliação para a Aprendizagem Organizacional Avaliação Participativa/ Colaborativa	Avaliação Responsiva; Avaliação Construtivista Responsiva; Avaliação Independente dos Objetivos (<i>Goal-free Evaluation</i>)	Avaliação Transformativa; Avaliação Deliberativa Democrática; Avaliação Feminista
Objetivos	Avaliação centra-se na explicação das causas, na aferição dos efeitos e em generalizações sobre a eficácia do tipo de intervenção.	Impactar a tomada de decisão, a mudança organizacional e os entendimentos conceituais das intervenções Manter decisores e outros <i>stakeholders</i> informados sobre os <i>inputs, outputs</i> e efeitos das intervenções	Descrição das necessidades, valores e perspetivas dos atores envolvidos na intervenção	Revelar pontos de vistas dos grupos marginalizados, interrogando sistematicamente as estruturas de poder

	Positivista	Pragmática	Construtivista	Transformativa
Metodologia	Quantitativa (primeiramente) Método científico Hipóteses	Correspondência de métodos com questões e finalidades específicas da avaliação Métodos mistos	Qualitativa (primeiramente) Métodos participativos	Métodos qualitativos (dialógicos), mas também quantitativos e mistos; Fatores contextuais e históricos são descritos, especialmente os que se referem à opressão
Orientação principal	Top-down	Top-down/Bottom-up	Bottom-up	Bottom-up
Utilização	Instrumental	Concetual e Instrumental	Cognitiva Empowerment	Promoção dos direitos humanos e justiça social
Vantagens	Objetividade; Capacidade de generalização; Fácil de legitimar.	Forte utilidade; Enfoque nas necessidades de informação dos <i>stakeholders</i> (sobretudo decisores); Potencia sentimento de pertença e de apropriação pelas pessoas envolvidas na intervenção	Sensibilidade a múltiplos pontos de vista; Construção de um modelo pluralista de percepções sobre os resultados; Serve de aprendizagem para todos os <i>stakeholders</i>	Forte orientação para a justiça social; Mobilização no sentido de tornar a avaliação justa e democrática; Desenvolvimento de competências culturais por parte de quem avalia
Desvantagens	Falta de aplicabilidade “no mundo real” (dificuldade de estabelecer controlo das variáveis); Falta de sensibilidade para os aspetos qualitativos e para a complexidade das relações humanas; Reduz influências complexas a simples relações de causalidade.	Risco das questões de avaliação serem dominadas pelos interesses de quem toma decisões; Dificuldade em sustentar o envolvimento dos <i>stakeholders</i> .	Problemas de viabilidade decorrentes da natureza aberta exploratória e participativa do processo de avaliação; Relutância em estabelecer prioridades ou em simplificar decisões para a decisão	Forte componente política – criticismo relacionado com a objetividade de quem avalia e dos resultados da avaliação; Pouca credibilidade junto de alguns grupos de <i>stakeholders</i> (especialmente entidades financiadoras); Dificuldades de aplicação prática

Fonte: Elaboração própria, a partir de, em especial: Guerra, 2002; Shaw, Greene e Mark, 2006; Stockmann, 2011; Mertens e Wilson, 2012; Stufflemeam e Coryn, 2014; Alkin, 2013.

Nos pontos que se seguem, estruturados a partir dos quatro paradigmas em que neste trabalho divido o “mundo” da avaliação – experimental, construtivista,

transformativo e pragmático – organizo as principais influências que afetaram a sua evolução e descrevo alguns dos modelos e abordagens que inspiraram.

1.2.1. Paradigma *experimental*: abordagens clássicas

Nos primeiros anos da pesquisa em avaliação, era considerado relevante o rigor metodológico determinado pelo positivismo. Neste paradigma, os modelos **experimentais** prevalecem, em ordem a revelar “as verdadeiras relações entre as forças causais” (Cook e Matt, 1990, *apud* Stockmann, 2011: 31), para que as decisões sobre as intervenções não “se baseiem em concepções falsas sobre a execução de um determinado programa”. Durante aquele que Shadish, Cook e Leviton (1991) classificam de *primeiro estágio das teorias sobre avaliação*, a avaliação empenhou-se no desenvolvimento de teorias focadas “na busca da verdade sobre as soluções efetivas para a resolução de problemas sociais” (*ibidem*: 67).

De acordo com Shadish, Cook e Leviton (1991), as teorias que representam o primeiro estágio da avaliação enfatizam o uso de procedimentos para serem validadas, respostas imparciais às questões sobre o desempenho das intervenções. Os também denominados *realistas científicos*, defendem a possibilidade de uma realidade independente capaz de ser descrita objetivamente.⁴ A avaliação clássica, racional ou científica tende a privilegiar estudos de natureza experimental ou quase-experimental, inspirados pelo trabalho de teóricos como Donald Campbell, Julian Stanley, Lee Cronbach e Richard Snow (Alkin, 2013; Mertens e Wilson, 2012) e que têm como propósito essencial o de determinar relações causais entre variáveis dependentes e independentes específicas. Tal exige a identificação de variáveis que reflitam os resultados das intervenções e a condição em que se encontram antes e depois da intervenção e/ou em comparação com a situação de um grupo de controle não abrangido pela intervenção. Estas abordagens têm

⁴ Discutindo os axiomas básicos deste paradigma, Christie e Alkin (2013: 17) dão conta da transição das visões da ciência durante o século XX do positivismo para o pós-positivismo. Enquanto os positivistas acreditam que o objetivo da ciência é descobrir a verdade, os pós-positivistas acreditam que o seu objetivo é tentar medir a verdade, embora esse objetivo não possa ser atingido porque toda a observação é falível e contém erros. Este tipo de realismo é designado “realismo crítico”, no qual o ideal de “aproximação” à verdade se estende do mesmo modo às visões de causalidade, uma vez que se acredita que as causas são observáveis e que ao longo do tempo podem ser estabelecidos prognósticos, ainda que persista sempre um grau de dúvida associado à conclusão. “Os valores e enviesamentos são anotados e considerados, no entanto, acredita-se que podem ser controlados no contexto da investigação científica” (*ibidem*: 17).

sido alvo de muitas críticas porque, além de serem inviáveis numa grande diversidade de contextos (Stufflebeam e Coryn, 2014: 147), não consideram de forma adequada nem os mecanismos que ligam as intervenções aos resultados nem a riqueza dos contextos heterogêneos (Henriques, 2006: 44).

Um marco importante da trajetória evolutiva da avaliação são os anos 50, nos quais surge a avaliação por objetivos, que constitui o paradigma dominante até aos nossos dias. A abordagem da **avaliação por objetivos**, da qual foi pioneiro Ralph Tyler (1949), enquadra-se nas premissas teóricas nas quais se fundamenta o modelo clássico da avaliação. Assume que as intervenções são desenhadas com a finalidade de alcançar objetivos específicos e que o êxito pode ser determinado em função da medida em que se alcançam os previstos. O processo deste modelo centrado nos objetivos, ainda que requeira condições exigentes, é relativamente linear e simples: especificação de metas e objetivos do programa; definição rigorosa de objetivos dispostos em hierarquia e em termos mensuráveis; e seleção dos instrumentos adequados para medir a consecução dos objetivos do programa (Stufflebeam e Coryn, 2014: 135). Esta perspetiva de avaliação, de que Tyler foi precursor, é enquadrada por Guba e Lincoln na **segunda geração da avaliação**, que os autores nomeiam de **fase da “descrição”**, e na qual a medição dos fenómenos é complementada pela sua descrição. Ainda que continue a ser uma abordagem privilegiada por quem faz avaliação, são muitas as críticas que lhe são formuladas. Alguns/mas autores/as (Ballart, 1992; Cuesta e Martínez, 2010) ressaltam, desde logo, que os objetivos das intervenções são frequentemente gerais, abstratos e confusos, sendo que, em muitos casos, é o trabalho da própria avaliação clarificar os objetivos e requisitos de determinada intervenção e verificar se os resultados correspondem ao esperado. Depois, a informação disponibilizada por este tipo de abordagem é muitas vezes demasiado estrita para permitir concluir sobre o valor de uma intervenção (Stufflebeam e Coryn, 2014: 135).

Tanto na fase da “medição” como na fase da “descrição” (orientada para objetivos), o papel de quem avalia é limitado ao de quem observa de forma neutra, que colige dados de uma forma científica. Autores tão representativos como Carol Weiss (2013) e Evert Vedung (2004), ambos a partir de uma perspetiva pós-positivista, reconhecem a inadequação e a inutilidade de utilizar desenhos experimentais e quase-experimentais na

grande maioria das situações e contextos de avaliação, a que se poderá acrescentar o seu questionamento ético. Não obstante, como nota Maria Bustelo (1998), a teoria da avaliação continua a atribuir-lhes um peso importante. As razões para o lugar central deste tipo de abordagem na teoria e manuais de avaliação podem ser encontradas, não apenas na dívida que a avaliação como disciplina ainda pode ter à investigação social, mas também “do próprio momento da revolução científica, nos termos kuhnianos, em que os paradigmas emergentes, como é o construtivismo emergente, coexistem, todavia, pouco ‘fluidos’, com outros paradigmas dominantes ao longo de muito tempo (Cuesta e Martínez, 2010: 86).

Numa crítica à procura de um conhecimento objetivo pelo *realismo científico*, Guba e Lincoln (1989: 7) afirmam que “abordar cientificamente a avaliação é perder completamente o seu carácter social, político e valorativo”. Estes autores denunciam as avaliações tradicionais por tenderem a privilegiar quem encomenda a avaliação, que é quem acaba por determinar as questões da avaliação, a metodologia a ser utilizada e os resultados a serem publicados. Este viés é, de certa forma, camuflado pelo uso de metodologias quantitativas e quase experimentais, baseadas no paradigma positivista.

O facto de a orientação positivista ter vindo a demonstrar, ao longo do tempo, limitações significativas na sua aplicação, nomeadamente ao nível da utilização dos seus resultados no processo de tomada de decisões políticas, tem levado alguns/mas estudiosos/as a privilegiarem abordagens políticas, com ênfase no emprego dos resultados da avaliação na tomada de decisão e, em geral, baseando-se em métodos de pesquisa qualitativos. Alinhado com esta perspetiva, Ballart conclui que “os projetos de investigação clássicos só podem ser concretizados em condições e circunstâncias muito limitadas. A exigência de uma avaliação de programas que não podem ser avaliados desta forma irá, pois, ser coberta pela mobilização de formas alternativas de avaliação” (1992: 149).

1.2.2. Paradigma *construtivista*: abordagens naturalistas

A avaliação naturalista representa uma alternativa ao paradigma clássico, reconhecendo a **pluralidade de sistemas de valores** que coexistem na sociedade. Este tipo de abordagem é acomodado pelos axiomas básicos do *paradigma construtivista*, que advoga que não existe uma realidade única, mas múltiplas realidades que emergem do

sistema de crenças subjetivo de cada indivíduo. Crê-se que as múltiplas realidades são subjetivas e variarão de acordo com “quem a observa”. No processo de pesquisa “quem observa” e “quem é observado” estão interrelacionados, e para determinar o que é “observado”, quem “observa” constrói uma realidade que se baseia e fundamenta no contexto e na sua experiência (Christie e Alkin, 2013: 17). Isto difere da epistemologia positivista, que propõe que “quem observa” e “quem é observado” podem ser independentes. Para os e as construtivistas, a pesquisa é balizada por, e não isenta de, valores e por isso os vieses devem ser objeto de conhecimento ao invés de os controlar. A causa e o efeito não são possíveis de distinguir uma vez que as relações são interdependentes e, por isso, simultaneamente, tudo tem influência em tudo. Os discípulos desta abordagem evocam o **método qualitativo** para conseguir compreender os vários pontos de vista e interpretações da realidade.

De acordo com a posição epistemológica do construtivismo, por princípio, não é possível fazer julgamentos acerca da verdadeira natureza do mundo: “eles mostram tão-somente se determinada interpretação é compatível com a natureza do mundo, se lhe serve ou não, mas não se é verdadeira (no sentido de ser a única versão correta) ” (Meinefeld, 1995, *apud* Stockmann, 2011: 33). Por isso, também na avaliação, esta atitude reivindica um procedimento científico que é diferente do invocado pela ciência empírica analiticamente orientada para o estudo das leis que presidem aos fenómenos naturais (vejam-se Lincoln e Guba, 2013; Patton, 2013, entre outros). Enquanto a visão “realista” acentua a primazia do julgamento independente por cientistas avaliadores/as, a visão “construtivista” vê quem avalia como facilitador e negociador de pontos de vista diferentes.

Em contraposição à abordagem tradicional, os autores consideram que a **avaliação é um processo político e social, orientado para valores**, e apresentam uma alternativa segundo a qual as reivindicações, interesses e questões dos diversos *stakeholders* (ou grupos de interesse) são os focos que organizam a avaliação. O paradigma construtivista, para o qual não há uma realidade dada como única, permite acomodar a diversidade de valores e interesses preconizada por uma abordagem alternativa, que Guba e Lincoln designam de avaliação construtivista (*responsive constructivist evaluation*):

É nossa intenção definir uma abordagem de avaliação emergente, mas madura, que vai além de uma ciência – que apenas acomoda factos –, para incluir uma infinidade de elementos humanos, políticos, sociais, culturais e contextuais implicados. Chamamos essa nova abordagem A Quarta Geração de Avaliação por sinalizar a nossa interpretação de que esta forma transcende as gerações previamente existentes, caracterizadas como orientadas para mensuração, descrição e julgamento, caminhando para um nível novo, cuja chave dinâmica é a negociação (Guba e Lincoln, 1989: 8).

Enquanto nos modelos tradicionais os parâmetros e limites são definidos *a priori*, a avaliação compreensiva determina-os por intermédio de um processo interativo e de negociação que envolve grupos de interesse. O termo *stakeholders* designa organizações, grupos ou indivíduos que podem ser agentes, alvos ou beneficiários do processo avaliativo. Esses grupos são formados por pessoas com características comuns, que têm algum interesse no desempenho, no produto ou no impacto do objeto da avaliação - estão, de alguma maneira, envolvidos ou potencialmente afetados pelo serviço que está a ser avaliado e pelas eventuais consequências do processo avaliativo (Lincoln e Guba, 2013; Guba e Lincoln, 1989).

Guba e Lincoln (1989) propõem o processo hermenêutico dialético como caminho para a realização de uma avaliação com enfoque construtivista e compreensivo. É hermenêutico, porque tem carácter interpretativo, e dialético porque implica comparação e contraste de diferentes pontos de vista. Os autores afirmam que a proposta desse processo não é legitimar uma determinada construção ou acometer as fraquezas de outras, mas antes construir uma forma de conexão que leve à exploração mútua de todas as partes.

A avaliação é vista, nesta perspetiva, como um processo pelo qual quem avalia e diferentes *stakeholders*, em conjunto e de forma colaborativa, criam (ou procuram) uma construção consensualizada do valor de determinados processos. Isso não significa necessariamente produzir informações irrefutáveis (isto é, empiricamente verificáveis), embora isso possa ser um resultado colateral.

1.2.3. Paradigma *transformativo*: abordagens emancipatórias

O *paradigma transformativo* pode ser considerado uma extensão do paradigma construtivista. Enfatiza as diferentes posições de poder dos diversos grupos de *stakeholders*, o que pode levar à sobrevalorização de determinados interesses (veja-se Mertens, 2009; Mertens e Wilson, 2012). Partindo do pressuposto de que há mais do que

uma realidade e de que cada uma destas realidades tem diferentes influências nos contextos social, cultural e político dos programas e respetivas avaliações, identificar e tornar as pessoas conscientes das posições de quem está em desvantagem é visto como um dos objetivos essenciais da avaliação. Este paradigma é alicerçado numa forte **agenda de defesa da justiça social e de direitos humanos** que é especificamente orientada para os grupos marginalizados, como sejam mulheres ou minorias étnicas, pessoas com deficiência, pobres e outras pessoas de grupos culturais considerados não dominantes. Mertens (2009: 3) explica que “o *paradigma transformativo* surgiu em resposta aos indivíduos que foram empurrados para as margens da sociedade ao longo da história e que estão a encontrar um meio de levar as suas vozes ao mundo da pesquisa”. Como tal, essa abordagem envolve diretamente quem investiga ou quem avalia no trabalho com as comunidades, no sentido de desafiar o *status quo* e analisar as posições de poder e privilégio, como um meio de promover a justiça social e os direitos humanos. As avaliações feministas (Brisolara, Seigart, e SenGupta, 2014) ou as avaliações do impacto de género (CE, 2000) são emblemáticas deste paradigma.⁵

Tendo o *paradigma transformativo* muito em comum com o paradigma construtivista, nomeadamente o seu carácter participativo e crítico, o que em última análise os diferencia é o enfoque daquele nas dimensões de poder, cultura, privilégio e justiça social. A sobreposição significativa entre as posições paradigmáticas ilustra aquilo a que Guba e Lincoln (2005) se refeririam como a fluidez de posturas e a mudança de categorias, num contexto em que quem investiga e quem avalia enfrenta as limitações metodológicas de um paradigma único para captar as múltiplas vozes que reclamam referência. Provendo um quadro teórico que se baseia noutras posições filosoficamente consentâneas e em abordagens qualitativas e de métodos mistos, o *paradigma emancipatório* parece fornecer uma alternativa robusta para apoiar a análise da complexidade do mundo.

⁵ Em Portugal este tipo de análise ainda se encontra numa fase incipiente, e tem sido particularmente tributário de trabalhos coordenados por Virgínia Ferreira no âmbito da avaliação da integração da perspectiva da igualdade entre mulheres e homens nos Fundos Estruturais Europeus (Ferreira *et al.*, 2007; Ferreira *et al.*, 2013).

As abordagens construtivistas à avaliação têm, porém, sido alvo de diversas críticas, sistematizadas no trabalho de Stufflebeam e Coryn (2014: 191-192). Um perigo recorrentemente apontado às abordagens construtivistas à avaliação é que, estando tão concentradas em servir uma missão social, possam não responder aos requisitos de avaliações sólidas e fiáveis. Outra crítica frequentemente enunciada é a de que, ao confiar aos *stakeholders* a autoridade para decidir sobre elementos chave da avaliação, sobretudo ao nível da interpretação dos resultados, quem faz avaliação capacita estas pessoas para usar a avaliação em seu proveito. Tal delegação de autoridade sobre assuntos importantes da avaliação pode torná-la vulnerável a enviesamentos e ao mau uso dos seus resultados. Os pontos fortes destas abordagens – de envolvimento de todas as partes interessadas e de forte orientação para princípios democráticos de equidade e justiça – potenciam também algumas dificuldades na sua aplicação. Os problemas de viabilidade são comuns à generalidade destas abordagens, uma vez associados, por exemplo, à dificuldade em sustentar a participação plena de todos os *stakeholders* ao longo do processo de avaliação, às dificuldades de contratação em processos de avaliação planeados em termos vagos, e à insistência de muitos clientes na obtenção de conclusões fundamentais sobre o mérito e valor das intervenções (*ibidem*: 207).

1.2.4. Paradigma *pragmático*: abordagens utilitaristas

No âmbito da discussão epistemológica em avaliação Christie e Alkin (2013), Mertens (2010) e Mertens e Wilson (2012) consideram ainda um terceiro (ou quarto) paradigma, relevante para compreender as influências teóricas do pensamento sobre avaliação: o *pragmático*. Quem segue esta linha adota a objetividade e subjetividade como duas posições num contínuo e consideram que as lógicas indutiva e dedutiva devem ser usadas em conjunto. Não obstante, rejeitam abarcar da mesma forma os axiomas dos paradigmas pós positivista e construtivista. Por exemplo, os/as pragmatistas assemelham-se mais aos pós positivistas relativamente às noções acerca da realidade externa, com o entendimento de que não existe nenhuma “verdade” absoluta acerca da realidade. Mais em linha com o pensamento construtivista, não obstante, os/as pragmatistas sustentam que há múltiplas explicações da realidade e que em cada momento existe uma explicação

que faz mais sentido. Por outras palavras, em determinado momento, uma explicação da realidade poderá ser considerada mais pertinente que outra. Além disso, pragmatistas, tal como pós-positivistas, acreditam que as causas podem estar relacionadas com os efeitos, ainda que temperem este pensamento com a advertência de que a certeza absoluta sobre a causalidade é impossível. O pragmatismo assemelha-se ao construtivismo na descrença de que a investigação seja isenta de valores; antes, considera-os importantes para o processo de investigação.

Este paradigma parece ser inspirador das abordagens da avaliação que preconizam o **uso instrumental dos seus resultados**, como seja a abordagem da **avaliação focada na utilização** (*Utilization-Focused Evaluation*) de Michael Patton (1997, 2008, 2013). O ponto de partida desta abordagem é de que as avaliações devem ser julgadas pela sua utilidade e uso efetivo; pelo que “o avaliador deve ser facilitador do processo de avaliação e desenhá-la tendo sempre em consideração que tudo o que é feito, do início ao fim, afeta o seu uso”, dizendo respeito “à forma como as pessoas reais no mundo real experienciam a avaliação e aplicam os seus resultados”, sendo o enfoque deste modelo no “uso pretendido e nos utilizadores pretendidos” (Michael Patton, 2013: 293). Reconhecendo que nenhuma avaliação é isenta de valores, o autor perfilha um modelo de avaliação que clarifique de quem são os valores que enquadram a avaliação, trabalhando com um grupo bem identificado (mas não necessariamente representativo) de *stakeholders*, que são responsáveis pela aplicação dos seus resultados e das suas recomendações. A avaliação focada na utilização não advoga nenhum conteúdo, método ou modelo de avaliação específico, nem nenhum tipo de utilização particular; é antes um “processo de apoio aos utilizadores prioritários na seleção dos conteúdos, modelos, métodos, teorias e usos mais adequados à sua situação particular” (*ibidem*: 293).

Um outro tipo de abordagem que poderá ser enquadrada no pensamento pragmatista, por ser desenhada a partir de um amplo espectro de conceitos e métodos de avaliação, respeita às formas **de avaliação participativas/colaborativas**. As formas participativas de avaliação emergiram em grande medida de processos de investigação-ação e a sua ênfase é no envolvimento de *stakeholders* no processo de avaliação. Cousins, um dos mais influentes teóricos desta abordagem, defende que o “locus do julgamento

sobre o mérito, valor e significância de um programa deverá caber a *stakeholders* localizados na comunidade onde decorre o programa” (Cousins, 2013: 346). O papel essencial de quem avalia é, neste sentido, promover o julgamento informado por parte dos *stakeholders*. Apesar de promoverem os juízos, as formas participativas de avaliação tendem a enfatizar sobretudo a melhoria das intervenções, mais do que os julgamentos sumativos sobre o mérito ou valor das intervenções (Stufflebeam e Coryn, 2014: 221).

A avaliação participativa promove a melhoria facilitando o **uso do processo** (mudanças no pensamento e comportamento e mudanças programáticas ou organizacionais nos procedimentos e na cultura que ocorrem entre quem está envolvido na avaliação, como resultado da aprendizagem que ocorre durante o processo); o **uso concetual** (no qual não tem lugar nenhuma ação direta baseada na avaliação mas o entendimento da intervenção é afetado); e o **uso simbólico** (em que a mera existência da avaliação, mais do que qualquer aspeto dos seus resultados, é usada para persuadir e convencer), mais do que o **uso instrumental** (envolvendo instâncias em que o conhecimento de avaliação é usado diretamente) (Stufflebeam e Coryn, 2014: 221, sublinhado meu).

Tal como a abordagem focada na utilização de Patton, a avaliação participativa é metodologicamente pluralista e flexível, mas, enquanto o foco central da primeira abordagem é responder com eficácia aos usos pretendidos para a avaliação por um grupo selecionado de utilizadores visados, a segunda não se restringe a servir um grupo de *stakeholders* predefinido. As maiores dificuldades de ambas as abordagens relacionam-se com a sustentação do envolvimento dos *stakeholders* e assegurar que os *stakeholders* empoderados não comprometam o rigor da avaliação (Stufflebeam e Coryn, 2014: 224). Stufflebeam e Coryn (2014) designam estas abordagens de *ecléticas*, por utilizarem de forma seletiva conceitos e métodos de amplo leque de abordagens e modelos de avaliação, no sentido de responder a necessidades de grupos particulares de *stakeholders*. Quem avalia e segue uma abordagem eclética poderá utilizar, em diferentes ocasiões, estudos de caso, desenhos experimentais, uma avaliação “sensível”, estudos baseados em objetivos ou uma avaliação orientada para a decisão.

A orientação pragmática enforma também, pelo menos em parte, algumas das perspetivas de avaliação que em diversos trabalhos (Stockmann, 2011; Christie e Alkin, 2013; Stufflebeam e Coryn, 2014) surgem classificadas como **abordagens centradas na melhoria e tomada de decisão**. Neste tipo de abordagens, as avaliações são desenhadas especialmente para apoiar a tomada de decisão por parte de *stakeholders* chave das intervenções (frequentemente, aqueles que encomendam as avaliações). As teorias com

esta orientação preocupam-se com o desenho de avaliações que não apenas informem a tomada de decisão mas que também tenham impacto na mudança organizacional e no entendimento concetual da intervenção (Christie e Alkin, 2013: 41). O modelo CIPP, desenvolvido por Daniel Stufflebeam (2014: 309-340, 2013: 243-260) é um dos modelos mais conhecidos, e constitui um acrónimo para quatro tipos de avaliação: contexto, *input*, processo e produto. A avaliação de *contexto* envolve a identificação de necessidades para decidir sobre os objetivos das intervenções. A avaliação de *input* avalia diferentes estratégias, planos de trabalho e orçamentos em alternativa. A avaliação de *processo* consiste na identificação de deficiências numa intervenção que está a decorrer no sentido de redefinir a implementação. A avaliação de *produtos* mede os resultados para informar as decisões sobre a continuidade ou enfoque das intervenções. Stufflebeam descreve a avaliação como um processo cíclico, que produz um fluxo contínuo de informações para a tomada de decisão no sentido de assegurar que as intervenções melhoram continuamente. Em suma, a sua abordagem envolve os *stakeholders* (normalmente em posições de decisão) no direcionamento da avaliação e na garantia de que responde às suas questões mais importantes; disponibiliza informação atempada e relevante para apoiar a tomada de decisão; e produz registos para prestação de contas. Apesar de neste tipo de avaliação o objetivo de “melhoria” prevalecer sobre o propósito de “prova”, ao procurar estabelecer uma avaliação definitiva do mérito e valor de uma intervenção, este modelo assume-se como objetivista e, por isso, também próximo da orientação positivista.

1.3. Abordagens managerialistas à avaliação

Fora do mundo da ciência, a avaliação é utilizada em organizações públicas, privadas e do terceiro setor para otimizar o seu planeamento e gestão, para analisar a prossecução dos seus objetivos e para averiguar sobre a efetividade e sustentabilidade dos serviços, programas, medidas e projetos que implementam. Podem ser fornecidas evidências do nível de sucesso com que as estratégias e políticas estão a ser implementadas.

No setor de mercado, onde a concorrência e a obtenção do lucro são predominantes, a avaliação, enquanto conceito radicado nas ciências sociais, é pouco

usada como ferramenta de gestão e garantia de qualidade (Stockmann, 2011). Estas organizações recorrem sobretudo a conceitos e instrumentos que têm a sua origem na “**gestão empresarial**”, alguns dos quais servem propósitos semelhantes e complementares aos da avaliação. Por essa razão, eles vêm sendo cada vez mais conhecidos e utilizados por organizações públicas e do terceiro setor na prossecução dos seus objetivos. **Este tipo de abordagem poderá ser vista como uma simplificação das perspectivas e modelos encontrados na teoria sobre avaliação (de programas).** Em termos dos seus fundamentos filosóficos a respeito da natureza da realidade, do conhecimento, dos princípios que inspiram e governam o processo de construção do conhecimento e dos instrumentos usados, as **perspetivas managerialistas de avaliação situar-se-ão entre a orientação positivista e a orientação pragmática.** Aproxima-se da orientação positivista ao tender a assumir uma observação desligada, neutra que pode recolher dados objetivos e imparciais, e ao privilegiar métodos e técnicas quantitativas. Por seu turno, aproxima-se da lógica pragmática, pela sua orientação utilitarista e foco na tomada de decisão e melhoria, e pelo uso do senso comum e pensamento prático.

Em anteriores análises sobre a situação das organizações do terceiro setor ao nível da avaliação (Tassie, Murray e Cut, 1998; Murray, 2005; Carman, 2007; Alaimo, 2008), verificou-se que a avaliação pode ocorrer de uma forma formal, sistemática, através da aplicação de um plano de avaliação concebido e implementado profissionalmente, ou pode ser desenvolvida com diferentes graus de informalidade, desde a recolha de dados para preparação dos relatórios até apreciações vagas sobre como decorrem as atividades. É precisamente esta perspetiva ampla de avaliação que tem transformado o campo da avaliação, com as pessoas das organizações a acumularem uma série de atividades de gestão que se associam à monitorização⁶ e medição do desempenho⁷. Essas atividades de gestão incluem, entre outras iniciativas, os sistemas de gestão da qualidade (revisão das

⁶ A monitorização é definida neste estudo como “sistema de registo de dados ou acompanhamento que recolhe, agrupa e analisa informação sobre a forma como as atividades estão a decorrer” (Schiefer *et al.*, 2006: 141). Implica o exame exaustivo e regular da utilização dos recursos, das realizações e dos resultados de uma intervenção e tem uma vocação de retroação sobre o conjunto dos procedimentos de gestão (Glossário CE, 2003).

⁷ A medição do desempenho é aqui entendida como “recolha sistemática e organizada de dados quantitativos a partir de um conjunto de indicadores-chave do desempenho da organização (ou programa)” (Poister, Aristigueta e Hall, 2014: 7).

práticas e procedimentos organizacionais no sentido de garantir que cumprem determinados requisitos de qualidade), as auditorias (verificação, por agentes independentes, da legalidade e da regularidade da utilização dos recursos em função de critérios e normas gerais), a avaliação de desempenho do pessoal, e o reporte (preparação de relatórios para financiadores e outros *stakeholders* que descrevem como é que o fundos concedidos foram empregues) (Cutt e Murray, 2000; Carman e Fredericks, 2008; Stockmann, 2011).

É a combinação de lógicas de que a avaliação é devedora – científico-metodológica, profissional e institucional – que torna a avaliação complexa. Mas, dada a dificuldade de compreensão do seu caráter híbrido, tende-se a identificá-la com instrumentos de apreensão mais fácil, como controlo, investigação e auditoria.

Por serem instrumentos usados amiúde no contexto da gestão empresarial, da gestão pública e da gestão do terceiro setor – frequentemente em substituição, mas também em complemento da avaliação – serão aqui brevemente introduzidas as bases conceituais do controlo e da auditoria.

1.3.1. Avaliação versus controlo e auditoria

O conceito de **controlo** desenvolveu-se a partir do campo da contabilidade, podendo ser entendido como o subsistema de gestão que fornece informações e dados relevantes sobre as mudanças que já ocorreram ou se espera vir a ocorrer no ambiente⁸, a fim de criar condições adequadas para ações de adaptação e medidas que possam influenciar o ambiente. O controlo fornece dados organizacionais internos para que quem decide tenha informação sobre a relação entre desenvolvimentos previstos e concretizados (comparação de metas e realizações), e que pondere ações corretivas (decisões de gestão).

⁸ A generalidade dos conceitos de controlo baseiam-se em considerações da teoria dos sistemas (veja-se, por exemplo, Horváth, 2011). A organização é vista como um sistema social que se conecta interactivamente com o seu ambiente. Constitui trabalho da gestão reagir adequadamente às exigências do ambiente e moldá-lo tanto quanto possível. A gestão deve avaliar de forma correta a complexidade e dinâmica do ambiente e desenvolver estratégias prospetivas que permitam salvaguardar a manutenção e prosperidade da organização. Quanto mais complexo o ambiente, maior a necessidade de controlo por si induzido. Para ser capaz de reagir prontamente e com flexibilidade às rápidas mudanças na procura e melhorar a sua capacidade de inovar, a gestão precisa de informação, cuja aquisição constitui uma das principais tarefas do controlo (Stockmann, 2008).

Pode também cumprir tarefas de coordenação em nome do sistema de gestão, criando as condições adequadas para a coordenação das ações dos diversos subsistemas individuais de gestão de uma organização por via da criação de sistemas de planeamento, controlo e informação, entre outros.

Se compararmos as tarefas de controlo com as de **monitorização**, encontramos algumas características comuns, mas também algumas diferenças importantes (Stockmann, 2011: 65-67). Tal como o controlo - mas ao contrário da avaliação - a monitorização é uma tarefa contínua, cujo objetivo é manter a gestão informada sobre a relação entre os desenvolvimentos previstos e alcançados (comparação de metas e realizações), utilizando dados de dentro e de fora da organização, de modo que a gestão possa fazer ajustamentos orientados para metas. Ambos os procedimentos têm a vocação de retroação sobre o conjunto dos procedimentos de gestão, alimentando a decisão com informações sem as avaliar, contribuindo assim para a tomada de decisão sem dela fazer parte. Existem, não obstante, diferenças importantes entre os dois procedimentos. Uma das diferenças é que enquanto a monitorização tem apenas uma função de disponibilização de informação, o controlo tem também uma função de coordenação. Outra diferença é que o espectro de informação de monitorização é muito mais vasto. Ao passo que os sistemas de controlo continuam, na prática, a concentrar-se sobretudo em aspetos financeiros, este é um domínio frequentemente negligenciado nos sistemas de monitorização, em benefício da informação sobre efeitos intencionais e não intencionais das intervenções. Por outro lado, ao passo que ao nível da monitorização as partes interessadas são frequentemente envolvidas - no estabelecimento de critérios e indicadores de avaliação, por exemplo -, ao nível do controlo esse envolvimento em geral não acontece - é determinado por "controles" fixos. O controlo segue uma abordagem de "cima para baixo", enquanto monitorização pode ser organizada como uma abordagem de "baixo para cima". Considerando que a monitorização é uma forma específica de avaliação contínua, a teoria da avaliação e os respetivos métodos também se lhe aplicarão. Estes métodos são bem mais extensos do que os procedimentos utilizados no controle, que se restringem em grande medida à contabilização dos custos. Enquanto o controlo é um procedimento bastante rígido, utilizando categorias de medição quase exclusivamente

monetárias e quantitativas, a monitorização é mais flexível, combinando dados qualitativos com dados quantitativos.

Uma vez que o **controlo** e a **monitorização** advêm de tradições distintas (respetivamente, da contabilidade e das ciências sociais) e são usados, por via de regra, em diferentes áreas (respetivamente, em organizações e em programas) não parecem ter sido desenvolvidas tentativas para unir os dois instrumentos (Stockmann, 2011: 67). Mas, tal como sustenta Reinhard Stockmann (*ibidem*), **unir estes procedimentos parece ser viável**. As abordagens da monitorização poderão contribuir para o enriquecimento das abordagens ao controlo, em termos dos seus conteúdos e metodologias, da sua “democratização” e abertura à realidade “subjativa” (isto é, às diversas perspetivas das partes interessadas). Poderão, por seu lado, beneficiar das abordagens ao controlo mediante a adoção de elementos e procedimentos associados à contabilização de custos.

Tal como o controlo, a **auditoria**⁹ desenvolveu-se a partir da contabilidade. Originalmente o seu principal objetivo era analisar dados financeiros de acordo com os princípios contabilísticos. Entretanto, o termo passou a ser aplicado a uma série de outras atividades. Além da auditoria financeira clássica, há agora formas de auditoria mais prospetivas, tais como a auditoria operacional e de gestão e a **auditoria de desempenho**:

Deve ser feita uma distinção entre a auditoria regular (de controlo, exemplo: controlo financeiro) que se centra sobretudo na conformidade com os procedimentos e regulamentos em vigor e a auditoria do desempenho que se interessa pela pertinência, a economia, a eficácia e a eficiência. (DAC, 2010)

As principais questões associadas à auditoria clássica (financeira) dizem respeito à integralidade, exatidão e transparência dos balanços financeiros. Enquanto a auditoria tradicional se centra exclusivamente no uso "correto" ou declaração "correta" do uso de recursos financeiros, o campo de procedimentos da **auditoria de desempenho** é muito mais vasto – envolve planeamento, trabalho de campo e reporte. Os seus procedimentos apresentam evidentes paralelos e semelhanças com algumas formas de avaliação. Também devido a esta proximidade de procedimentos, encontramos com facilidade literatura que

⁹ A auditoria é definida por diversos organismos internacionais nos seguintes termos: Atividade de controlo em matéria de qualidade, realizada de forma objetiva e independente e destinada a melhorar as operações de uma organização e a aumentar o seu valor. A auditoria ajuda uma organização a alcançar os seus objetivos graças a uma abordagem sistemática e rigorosa para apreciar e melhorar a eficácia da gestão de riscos, do controlo e dos processos de governação. (DAC, 2010)

compara avaliação e auditoria (veja-se, por exemplo, Pollitt e Summa, 1997; Bastoe, 2006), ao contrário do que sucede no caso do controlo.

Todavia, embora conceitualmente próxima da **avaliação**, com a qual partilha o objetivo de melhoria dos serviços ou programas, uma visão mais aproximada dos processos associados à auditoria de desempenho revela que os **pressupostos teóricos dos respetivos procedimentos, questões específicas e métodos utilizados na produção e análise de dados divergem consideravelmente** dos pressupostos, questões e métodos da avaliação. Desde logo, a auditoria e a avaliação têm as suas origens em tradições distintas. Enquanto a auditoria tem origens na contabilidade do século XIX, a avaliação não se desenvolve na sua forma moderna de ciência social antes dos anos de 1960. A auditoria está mais fortemente preocupada com questões da boa gestão, enquanto a avaliação vai mais longe e se preocupa sobretudo com os resultados obtidos e os impactos gerados, bem como com questões como a relevância, pertinência ou sustentabilidade das intervenções (SEAV, s/d). A avaliação, ao contrário da auditoria, está mais "preocupada com a teoria e com a explicação do que com a regularidade e conformidade" (Pollitt e Summa 1997: 88). A avaliação continua a ser mais uma ciência do que um procedimento administrativo, empregando uma ampla variedade de conceitos e métodos das ciências sociais, e desenvolvendo um extenso conjunto de tarefas. A auditoria, por oposição, é um conceito com uma formulação mais clara, compreendendo uma gama bem mais limitada de tarefas. Enquanto as avaliações podem contemplar *inputs*, *outputs* e resultados, investigar efeitos intencionais e não intencionais, de longo ou de curto prazo, ou a sustentabilidade, utilidade e relevância social das intervenções e as relações entre áreas específicas de impacto, as auditorias de desempenho concentram-se sobretudo na fase de implementação, nas realizações e resultados imediatos. Há também diferenças em relação aos destinatários dos relatórios: enquanto eles são claramente estipulados no caso das auditorias, em geral os públicos das avaliações são mais variados. Uma das principais diferenças entre auditoria e avaliação fica patente no tipo de questão que é colocada: "a auditoria está interessada em saber em que medida aquilo que é feito está em conformidade – a questão "o quê". Os avaliadores tentam compreender o que produz os efeitos desejados e indesejados – a questão "porquê" (Pollitt e Summa 1997: 89). A lógica da auditoria baseia-se na resposta a

questões normativas, por via da comparação do estado atual das intervenções com os critérios e normas estabelecidos e em que as relações causa-efeito tendem a ser obscurecidas. A avaliação é vista como uma abordagem mais ampla, que se ocupa de questões descritivas, normativas e de causa e efeito (Bastoe 1999: 104). Segundo alguns autores (Pollitt e Summa 1997; Stockmann, 2011), as principais diferenças resultam do contexto institucional da auditoria e da avaliação. De acordo com os autores, as entidades governamentais de auditoria têm um elevado grau de unidade e de solidariedade coletiva: "Elas acomodam uma comunidade profissional coerente, bem definida, apoiada por normas internacionalmente aceites e fortes poderes estatutários" (Pollitt e Summa 1997: 104). Já as entidades de avaliação têm poder institucional unicamente interno e não formam uma equipa homogénea; antes competem entre si por trabalhos de avaliação (Stockmann, 2011: 86). O que distingue avaliadores ou entidades de avaliação são os seus métodos de trabalho mais profundos, teoricamente fundamentados e metodologicamente imaginativos, as ideias criativas e originais e o elevado grau de competência que é tanto social como profissional (*ibidem*).

Apesar das diferenças que persistem, parece evidente que **nos últimos anos a auditoria se tem aproximado da avaliação**, designadamente em fontes de informação utilizadas. As semelhanças e a relação de complementaridade entre auditoria e avaliação fazem supor que existe um potencial considerável para a aprendizagem mútua.

Sejam **abordagens da contabilidade e gestão empresarial** (além do controlo e da auditoria, poderão destacar-se também, entre outras, o *balanced scorecard*¹⁰ e o *benchmarking*¹¹) ou **abordagens influenciadas pela investigação em ciências sociais**

¹⁰ O *Balanced Scorecard* é um instrumento de gestão que visa a construção de um conjunto de indicadores quantitativos que permitem avaliar a capacidade da organização para cumprir a sua missão e objetivos estratégicos. Estes indicadores estruturam-se em torno de 4 perspetivas: inovação e liderança; processos internos; clientes e gestão financeira. Os indicadores de cada uma das perspetivas estabelecem entre si relações de causa-efeito, as quais são baseadas em hipóteses que permanentemente têm de ser revistas (Kaplan e Norton, 1996).

¹¹ Padrões de natureza qualitativa e quantitativa para efeitos de comparação do desempenho de uma intervenção. Este modelo padrão será, muitas vezes, o melhor no mesmo domínio de intervenção ou num domínio relacionado. A avaliação comparativa de desempenho é aplicada, a nível nacional ou regional, quando existe informação comparativa entre práticas consideradas boas e menos boas. O termo "benchmarking" também é usado para se referir à comparação de condições contextuais entre territórios (CE, 2003). Segundo a definição do EuroDicAutom (base terminológica da UE: <http://iate.europa.eu/iatediff/SearchByQueryLoad.do?method=load>): Instrumento de melhoria ou

(monitorização, avaliação e, até certo ponto, auditoria), em qualquer dos casos **elas geram informação relevante para proporcionar uma base racional para a tomada de decisões de gestão**, podendo, assim, ser também entendidas como *instrumentos de gestão*. Com base no trabalho de Stockmann (2011: 87-91), será aqui feita uma apresentação sintética e necessariamente simplista das **semelhanças e diferenças entre os conceitos originários da gestão empresarial e as abordagens da monitorização e da avaliação, com raiz na pesquisa em ciências sociais**. Será necessariamente uma comparação simplista porque para além de não existir um único conceito de controlo ou de auditoria, cada um deles tem inúmeras *nuances* e distintivos e vai sofrendo uma série de desenvolvimentos ao longo do tempo. Este exercício de comparação pretende oferecer uma visão genérica que torne possível alocar determinados procedimentos em termos da sua relação com a monitorização e avaliação.

Como se pode ver no quadro abaixo, os conceitos oriundos da gestão empresarial têm uma série de semelhanças entre si que os distinguem, enquanto grupo, das abordagens da pesquisa em ciências sociais.

aperfeiçoamento funcional ou estrutural de uma organização ou sistema baseado na avaliação contínua e sistemática de produtos, serviços, práticas ou métodos feitos por comparação com elementos similares, considerados de excelência e tomados como padrão ou referência. Este processo, concebido e vulgarizado no contexto empresarial como uma técnica de gestão, mas que atualmente se aplica a contextos mais alargados, é habitualmente classificado em três tipos: “aferição interna”, em que os elementos de comparação se encontram no interior de uma mesma organização ou empresa; “aferição competitiva” ou “concorrencial”, em que os marcos de referência se situam no mesmo setor de atividade numa perspetiva de concorrência; “aferição funcional”, também designada “genérica” ou “universal”, em que a comparação se faz entre elementos funcionalmente equivalentes independentemente da área ou setor.

Quadro 3. Comparação entre instrumentos de gestão/avaliação

C = Controlo; BSC = Balanced scorecard; BM = Benchmarking; A = Auditoria; M = Monitorização; Aval = Avaliação.

Características		C	BSC	BM	A	M	Aval
Origem científica	- Gestão empresarial	x	x	x	x		
	- Ciências sociais					x	x
Uso principal	- Setor privado	x	x	x	x		
	- Setor público			x	x	x	x
Tarefas	- Monitorização	x	x	x	x	x	
	- Explicação						x
Foco de interesse	- Processos internos	x	x	x	x	x	
	- Processos externos					x	x
Método	- Quantitativo	x	x	x	x	x	x
	- Qualitativo					x	x
Orientação principal	- Top-down	x	x	x	x	x	x
	- Bottom-up					x	x
Padrão de apreciação	- Predeterminado	x	x	x	x	x	
	- Variável						x
Periodicidade	- <i>Contínuo</i>	x	x	x		x	
	- Esporádico				x		x
Agentes	- Internos	x	x	x		x	x
	- Externos				x		x
Área de investigação	- Restrita (financeira, clientes, ...)	x	x	x	x		
	- Abrangente					x	x
Lógica de investigação	- Normativa (metas/realizações)	x	x	x	x	x	
	- Orientada para a relação causa-efeito						x
Perspetiva de investigação	- prospetiva						x
	- Retrospetiva	x	x	x	x	x	x
Audiências	- Unidimensional (gestão, legislat.)	x	x	x	x	x	
	- Multidimensional						x
Crítérios de apreciação	- Efetividade		x	x	x	x	x
	- Eficiência	x	x	x	x	(x)	(x)
	- Eficácia					x	x
	- Sustentabilidade					x	x
	- Relevância social					x	x

Características		C	BSC	BM	A	M	Aval
Consideração interesses dos stakeholders	- Sim					×	×
	- Não	×	×	×	×		
Implicações dos resultados para a gestão	- Vinculativas	×	×	×	×		
	- Não vinculativas					×	×

Fonte: Stockmann, Reinhard (2011) *A practitioner handbook on evaluation*. Cheltenham: Edward Elgar, Quadro 3.7.

As **maiores diferenças** no quadro 3 podem ser encontradas entre a **avaliação**, por um lado, e as **abordagens da gestão empresarial**, por outro. Enquanto as últimas tendem a concentrar-se em processos internos das organizações, negligenciando os processos externos, com a avaliação acontece frequentemente o inverso. Os critérios associados à capacitação para o desempenho das organizações que implementam as intervenções raramente são usados em favor de uma particular atenção aos efeitos esperados e não esperados. Observação semelhante pode ser feita a respeito dos critérios de apreciação usados por estes dois tipos de abordagens: ambos os grupos de instrumentos conferem uma boa dose de atenção às questões da efetividade, isto é, ao alcance dos objetivos. Todavia, as abordagens com origem na gestão empresarial fundamentam as suas análises sobretudo em questões de eficiência, as quais tendem a ser descuradas pela avaliação. Em compensação, o âmbito das avaliações é bem mais abrangente e fundamentam as suas apreciações numa maior diversidade de critérios, como a relevância, a eficácia, e a sustentabilidade. Enquanto as primeiras são tendencialmente orientadas para resultados operacionais, as segundas são sobretudo orientadas para a compreensão das relações de causalidade e para os impactos.

Enquanto algumas destas abordagens foram desenvolvidas para responder sobretudo às necessidades das organizações privadas com fins lucrativos (controlo, *balanced scorecard*), outras tendem a ser mais usadas no setor público e na gestão de programas (monitorização, avaliação). O *benchmarking* e a auditoria não podem, todavia, ser classificados desta forma simples. No que toca à origem dos seus princípios e áreas de aplicação, em termos conceituais e metodológicos, enquadrar-se-ão mais nos instrumentos de gestão do setor privado. Todavia, o *benchmarking* e a auditoria são também muito usados no setor público. Uma análise de *benchmarking* tanto pode ser realizada com dados

obtidos a partir de uma auditoria como de uma avaliação. As auditorias são usadas tanto para certificação de empresas como para o escrutínio e controlo de programas públicos. Sendo assim cada um destes conceitos tem diferentes graus de afinidade com **modelos de gestão** tais como gestão da qualidade¹² ou a Nova Gestão Pública.

Embora alguns dos novos modelos de gestão, como seja a Nova Gestão pública (e, no seu enquadramento, a gestão por objetivos¹³ ou a gestão orientada para resultados¹⁴), reivindicuem dados sobre realizações e respetivos efeitos, a fim de tornar possível a gestão orientada para resultados, o potencial da avaliação é frequentemente menosprezado nas organizações em favor de abordagens focadas em indicadores operacionais e na eficiência (Barman, 2007) como sejam o controlo e o *benchmarking*. Esta restrição a fatores de gestão que se inspiram na gestão empresarial é, no entanto, cada vez mais criticado como "reducionismo económico" (Kuhlmann 2003: 12 *apud* Stockmann, 2011: 95).

Algumas destas abordagens oriundas da gestão empresarial servem sobretudo para monitorizar os processos internos das organizações, enquanto a avaliação possibilita a avaliação dos processos externos de mudança e a compreensão de como e porquê estas mudanças ocorrem. Particularmente no terceiro setor, que subordina os objetivos económicos à criação de valor social e ambiental, uma combinação de instrumentos com enfoque em processos operacionais (**controlo, *balanced scorecard***), com a recolha contínua de informação **quantitativa e qualitativa sobre realizações, resultados e impactos (monitorização)**, e com **avaliações**, que analisam não só a realização dos objetivos, mas também testam esses mesmos objetivos, comparam efeitos desejados e efeitos indesejados, utilizam critérios de análise como a sustentabilidade e a relevância e

¹² Método para assegurar que todas as atividades necessárias à conceção, desenvolvimento e implementação de um produto ou serviço são eficientes e eficazes, respeitando o sistema da qualidade (SEAV, s/d).

¹³ Tipo de gestão na qual os objetivos são definidos em "cascata", efetuando-se o seu desdobramento a partir do topo da organização, de uma forma sistemática e estruturada. Foi no livro "The Practice of Management" que em 1954 Peter Drucker introduz pela primeira vez o conceito de Gestão por Objetivos e os seus princípios: "Cascata" de objetivos e metas; objetivos específicos para cada membro da organização; tomada de decisão participada; prazos concretos; avaliação de desempenho e feedback.

¹⁴ Estratégia de gestão que se centra no desempenho, na obtenção de resultados, efeitos e impactos (DAC, 2010).

procuram as causas das mudanças observadas, pode ser a chave para uma gestão organizacional bem-sucedida.

1.3.2. Entre diferentes tradições disciplinares: Avaliação *versus* medição do desempenho

Uma confusão muito comum na literatura e na prática sobre a avaliação de projetos, serviços ou organizações reside na percepção da avaliação como sinónimo de medição de desempenho (Ellis e Gregory, 2008). A distinção entre estas duas abordagens é aqui trazida enquanto **arquétipo da distinção entre práticas alicerçadas na investigação em ciências sociais e as práticas managerialistas, alicerçadas na gestão empresarial.**

A avaliação, enquanto campo, tem vindo a transformar-se nas últimas duas décadas pelo movimento mais amplo que tem tido lugar nas organizações públicas e do terceiro setor no sentido de construírem e implementarem sistemas de medição do desempenho organizacional ou programático (McDavid, Huse e Hawthorn, 2013). A medição do desempenho, ao contrário da avaliação, tem origem no mundo empresarial, tendo sido os sistemas de medição do desempenho originalmente concebidos e criados a partir das perspetivas de empresas com fins lucrativos (Speckbacher, 2003).

A Organização para a Cooperação e Desenvolvimento Económico (OCDE) define medição do desempenho como “sistema que permite apreciar os desempenhos das intervenções relativamente aos objetivos estabelecidos”. Poister, Aristigueta e Hall (2014) avançam uma definição mais detalhada e concetual, afirmando-a como a “recolha sistemática e organizada de dados quantitativos a partir de um conjunto de indicadores-chave do desempenho da organização (ou programa)” (*ibidem*: 7). Estes/as autores/as distinguem *medição do desempenho* de *gestão do desempenho*, que definem como “o processo de definição, acompanhamento e utilização de indicadores objetivos de desempenho de organizações e programas para informar a gestão e tomada de decisão numa base regular” (*ibidem*: 1). Ou seja, a gestão do desempenho implica a utilização rotineira e estratégica das informações de desempenho por parte da gestão.

A gestão do desempenho, que se enquadra na estratégia de gestão orientada para os resultados, surgiu como uma abordagem de gestão organizacional que faz parte

do amplo movimento de Nova Gestão Pública (NGP)¹⁵. A Nova Gestão Pública (Hood, 1991) reorientou a gestão do setor público para uma maior eficiência, com destaque para a utilização de práticas de mercado emprestadas da gestão do setor empresarial. O resultado foi uma explosão do planeamento estratégico, que induz uma forte orientação para metas, assim como esforços para avaliar o desempenho e para ajustar a estratégia no sentido de a reforçar.¹⁶ O movimento reivindica uma maior discricionariedade em matéria de gestão em troca de uma maior responsabilização pelos resultados. Assim, embora a avaliação e a medição de desempenho se tornem componentes estratégicos da NGP, **o foco na prestação de contas e nos resultados levou a um papel mais proeminente da medição do desempenho e a um papel mais circunscrito da avaliação** na reforma do setor público (Mayne e Rist, 2006; McDavid e Huse e Hawthorn, 2013). Pode dizer-se que a “avaliação” é aqui reforçada, mas assumindo novas formas e sob novas expressões, designadamente, prestação de contas, avaliação da satisfação de clientes, garantia de qualidade, *benchmarking* e medição do desempenho.

A ênfase na medição e na comparação do desempenho estimulada pela NGP tem influenciado a adesão das organizações do setor público e do terceiro setor a **mecanismos de gestão da qualidade** (ISO, EQUASS, etc.). Embora possam ser encarados como

¹⁵ Desde os últimos anos vem sendo frequente a promoção de uma cultura de gestão nas administrações públicas, considerando-se que para a sua concretização é necessária uma determinação de objetivos clara e específica, e a promoção de uma cultura de resultados, centrada no cumprimento de objetivos estabelecidos e na análise dos seus resultados. A constatação dos problemas de eficiência na atuação pública fez com que a reformas na administração tenham colocado acento na atualização e melhoria nos modelos de gestão pública, apoiados na ideologia de certas correntes, em particular a Nova Gestão Pública (NGP). Na sequência de fortes modificações na economia mundial nos anos 70 e consequentes críticas à inadaptação do Estado Burocrático aos desafios de uma economia cada vez mais competitiva, este modelo, com origens no Reino Unido e Estados Unidos da América, emerge nos anos 80, ancorado nas correntes de modernização da gestão empresarial e apoia-se na aplicação de técnicas e modelos de mercado no campo administrativo (Rodrigues, 2011). Para resolver os problemas de uma burocracia ineficaz, esta corrente propõe a separação entre administração e política, aparecendo, com uma roupagem politicamente neutra. Esta visão, assente na filosofia da forte capacidade que a gestão empresarial tem para dar respostas aos problemas do mundo atual, é apresentada de certo modo como sendo assética e tecnocrata, tendendo a considerar a NGP como uma resposta instrumental (*ibidem*: 142).

¹⁶ É exatamente a introdução de medidas managerialistas, sob a influência da NGP, que levou diversos países da OCDE, incluindo Portugal, a desenvolverem esforços para implementar sistemas de avaliação de desempenho nas respetivas administrações públicas. A avaliação de desempenho é definida, no contexto da gestão orientada para os resultados, como um “sistema de avaliação de resultados de pessoas, das equipas e das organizações, em função dos objetivos previstos ou dos critérios/padrões pré-definidos” (SEAV, s/d).

complementares à avaliação, trata-se de processos distintos. Enquanto nos sistemas de garantia da qualidade, os padrões definidos constituem os critérios a partir dos quais os juízos são produzidos, na avaliação os critérios podem ser estabelecidos mediante termos de referência particulares, ou podem surgir através do processo de avaliação em si (Ellis e Gregory, 2008: 51).

Na esteira da NGP, e em nome de prestação de contas, eficiência e eficácia, a medição do desempenho disseminou-se através das fronteiras internacionais, nacionais e locais. Há quem chegue a identificar a medição de desempenho com uma **tendência histórica independente, ou “vaga” de avaliação** (Vedung, 2010). Isso reflete o forte impacto que a medição do desempenho tem enquanto uma forma de avaliação dominante (Bjørnholt e Larsen, 2014).

A medição de desempenho tornou-se uma prática comum no setor público e cada vez mais frequente no terceiro setor. Numa era de resultados, o desafio de convencer organizações, gestores/as, fazedores/as de políticas, entidades financiadoras, órgãos estatais e afins da importância de focar em resultados e de monitorizar o desempenho numa base contínua, sistemática, tem sido, em parte, conseguido. Ao longo dos últimos vinte e cinco anos, a medição de desempenho tem-se tornado amplamente aceite, focada nos resultados, mais sofisticada e institucionalizada (Poister, 2010), ainda que não isenta de críticas. A medição do desempenho é controversa entre estudiosos e profissionais da avaliação: há quem advogue que a profissão abarca a medição do desempenho (Feller, 2002), há quem seja cético e exclua a medição do desempenho do campo da avaliação, referindo-se-lhe como um instrumento ao serviço da gestão e não da avaliação (Bernstein, 1999).

Existem, de facto, **diferenças consideráveis nas conceções a respeito da articulação entre monitorização do desempenho e avaliação**, em particular sobre até que ponto estas abordagens estão desalinhasadas ou se complementam (Wholey, Hatry, Newcomer, 2010: 673-675; Hatry, 2013; McDavid, Huse e Hawthorn, 2013; Nielsen e Hunter, 2013). Os estudos de avaliação procuram fornecer informações sobre qualquer tipo de relação que pode ser estabelecida entre a intervenção de um programa ou organização e os respetivos resultados (Wholey, Hatry e Newcomer, 2010: 673). As

avaliações de impacto, em particular, procuram fornecer informações sobre os resultados da intervenção e atribuir explicações para a ocorrência desses resultados, principalmente a medida em que a intervenção contribuiu para os resultados. Essas avaliações podem requerer recursos substanciais. É por isso que, num determinado ano, as avaliações abrangem geralmente apenas uma pequena parte da intervenção da organização (*ibidem*: 673; Hatry, 2013: 24).

Como notam Poister, Aristigueta e Hall (2014: 27), entre profissionais de avaliação existe a noção de que a avaliação, particularmente a avaliação de programas, é mais sofisticada metodologicamente, mais definitiva nas suas conclusões, e mais útil para determinar o valor real de uma intervenção do que a de medição de desempenho. Estes pontos de vista são, no entender dos autores, pertinentes, já que a avaliação de programas tem-se apoiado, tradicionalmente, em projetos sofisticados de pesquisa, metodologias complexas e avaliações experimentais ou quase-experimentais rigorosas que ultrapassam o alcance da medição de desempenho. Esta, por seu turno, é um processo relativamente simples: os objetivos das intervenções e os resultados correspondentes são identificados, identificam-se medidas para apurar os resultados, e os dados são reunidos para permitir a quem gere e outras partes interessadas monitorizar o desempenho programático/organizacional (McDavid, Huse e Hawthorn, 2013: 4). Visa sobretudo estimar os resultados das intervenções, sem uma análise aprofundada dos efeitos das mesmas sobre esses resultados. Geralmente, as organizações que desenvolvem sistemas de medição de desempenho monitorizam grande parte da sua intervenção, e a informação sobre os resultados é disponibilizada com intervalos frequentes (mensal, trimestral, semestral ou anual) (Wholey, Hatry, Newcomer, 2010: 673). Os estudos de avaliação, por sua vez, são substancialmente mais informativos e fornecem muito mais orientações para as grandes decisões organizacionais e programáticas (Hatry, 2013; Nielsen e Hunter, 2013). A monitorização do desempenho tende a ser mais uma ferramenta de gestão, devido à frequência e, assim, assiduidade, das informações, e tem a vantagem de manter o foco nos resultados numa ampla variedade de áreas de intervenção (Wholey, Hatry, Newcomer, 2010: 674). Mas, uma vez que, em geral, a gestão desempenha um papel fundamental na

medição e elaboração de relatórios de desempenho, as visões mais críticas encaram a medição do desempenho como apenas um aspecto da gestão organizacional.

O quadro 4 fornece uma perspectiva muito resumida sobre as diferenças entre avaliação (de programas) e medição de desempenho e as principais vantagens e desvantagens dos dois tipos de abordagens.

Quadro 4. Comparação entre avaliação e medição de desempenho

Características	Avaliação	Medição do desempenho
Origem científica	Investigação em ciências sociais	Gestão empresarial
Frequência	Irregular, esporádico	Regular, contínua
Cobertura	Programas/projetos/serviços específicos	Abrange a maioria dos programas/atividades/serviços
Custos	Podem ser elevados	Tendem a ser distribuídos
Recursos técnicos	Competências especializadas (designadamente analíticas)	Menos exigentes em termos de competências específicas por ser um processo relativamente simples
Tipos de informação	Indicadores quantitativos e qualitativos individualizados <i>Inputs, outputs, resultados, impactos</i>	Indicadores quantitativos reproduzidos ao longo do tempo <i>Inputs, outputs, resultados</i>
Recolha de informação	Pesquisa documental, entrevistas, inquéritos, sistemas de informação	Pesquisa documental, entrevistas, inquéritos, sistemas de informação
Análise da informação	Triangulação de múltiplas fontes de informação A atribuição dos resultados da intervenção pode ser um objetivo em si mesmo	Sustentada por poucas fontes de informação Atribuição assumida
Qualidade da informação	Olhar próximo sobre os dados Seleção de indicadores de resultado de difícil mensuração.	Olhar distante sobre os dados Seleção de indicadores de resultados facilmente mensuráveis.
Julgamento avaliativo	Descritivo e prescritivo Procura razões para o mau/bom desempenho	Descritivo Apenas diz “o resultado”, não o porquê
Disponibilidade da informação	Em geral, o prazo para a conclusão de um estudo de avaliação é relativamente longo Fluxo de dados descontínuo	Informação sobre o desempenho “rápida” e atempada Um fluxo contínuo e atualizado de dados
Utilidade	Suporte às principais decisões	Melhoria contínua
Viabilidade nas pequenas organizações	Pequenas organizações têm maiores dificuldades em conduzir estudos de avaliação sofisticados	Exequível em pequenas organizações. O acompanhamento dos principais indicadores de desempenho tende a ser valorizado pela gestão da generalidade das organizações

Fonte: Elaboração da autora, com base em: Wholey, Hatry, Newcomer, 2010; Hatry, 2013; McDavid, Huse e Hawthorn, 2013; Nielsen e Hunter, 2013.

Apesar das diferenças, tal como é defendido em outros trabalhos (Wholey, Hatry e Newcomer, 2010; McDavid, Huse e Hawthorn, 2013; Hatry, 2013; Nielsen e Hunter, 2013; Poister e Aristigueta, 2014), **entendo que estes dois processos são complementares e podem ser integrados em aspetos importantes**. Na literatura emergem quatro tipos de complementaridade entre avaliação e gestão ou medição de impacto (Nielsen e Ejler, 2008: 173-174; Nielsen e Hunter, 2013: 120; Lahey e Nielsen, 2013: 51):

1. Complementaridade sequencial – As informações de monitorização podem gerar questões a ser respondidas por estudos de avaliação; e vice-versa: os estudos de avaliação podem gerar conhecimento que exige um acompanhamento contínuo do desempenho.
2. Complementaridade de informação – Tanto a medição de desempenho como a avaliação emergem das mesmas fontes de dados, evitando a construção de fluxos de informação em paralelo.
3. Complementaridade organizacional – A informação dos sistemas de medição do desempenho e de avaliação tende a ser acoplada, o que requer que as informações de monitorização e de avaliação sejam canalizadas através da mesma unidade administrativa.
4. Complementaridade de métodos – Ambas as formas de produção de conhecimento partilham processos e ferramentas semelhantes de planeamento, de obtenção de informação, de análise e aferição, e de conversão de dados em informações uteis.

McDavid, Huse e Hawthorn (2013) argumentam que a **medição de desempenho**, independentemente de quem a faz, **assenta na compreensão dos princípios e práticas da avaliação**. Ambos os processos dependem de um entendimento comum da teoria da intervenção da qual derivam medidas concetuais essenciais – isto é, da **lógica do processo** através do qual uma intervenção converte *inputs* e recursos em produtos e resultados através das suas diversas atividades. Por outro lado, muitas das **competências básicas** adquiridas por quem faz avaliação podem ser aplicadas à medição do desempenho. Do mesmo modo, o **pessoal da gestão e outros atores envolvidos** no desenvolvimento e implementação da medição de desempenho defrontam-se com **problemas similares** a quem avalia. Além de poder ser um **componente importante da avaliação**, a medição do desempenho tende a encorajar o desenvolvimento de estudos de avaliação.¹⁷

¹⁷ Tal ocorre de duas formas, no entender de Wholey, Hatry e Newcomer (2010: 674): em primeiro lugar, as informações sobre os resultados de um sistema de medição do desempenho levantam

Vejo, neste trabalho, a medição de desempenho como abordagem complementar à avaliação, mas não como sinónimo ou substituto das avaliações. Apesar de esta complementaridade ter vindo a ser reconhecida, em algumas instâncias, os esforços para abraçar a **medição de desempenho ofuscaram a avaliação** (Ellis e Gregory, 2008; Bradford e Chouinard, 2010; McDavid e Huse e Hawthorn, 2013). Johnsen (2013) observa que, em alguns casos, o acompanhamento e a avaliação não são meramente formas complementares de conhecimento, mas na realidade estão a competir por recursos escassos. As tensões na articulação entre avaliação e medição do desempenho são, aliás, reconhecidas por agências internacionais impulsionadoras da estratégia de gestão orientada para os resultados, a qual atribui à medição do desempenho um papel central (DAC, 2000)¹⁸. No entender de McDavid e Huse e Hawthorn (2013), existem cada vez mais evidências de que foram goradas as promessas de que a medição de desempenho se constituiria numa ferramenta de responsabilização e de gestão do desempenho. McDavid, Huse e Hawthorn não estão sozinhos nestas críticas. Entre os desafios importantes que se colocam à medição e gestão do desempenho contam-se críticas relevantes de proeminentes estudiosos e estudiosas da área. Moynihan (2008) argumenta que a prevalência do uso simbólico face ao emprego substancial das informações de desempenho torna o processo uma fachada. Segundo o autor, para que a medição de desempenho possa gerar impactos, ela deve ser usada substantivamente para guiar a ação. Piotrowski e Rosenbloom (2002) ressaltam a preocupação com os aspetos de desempenho que não têm sido medidos, incluindo os valores democráticos e constitucionais, tais como a **transparência, a equidade ou a justiça**. Radin (2006) faz eco destas preocupações,

frequentemente questões sobre as razões que sustentam os bons ou os maus resultados. Quando o que está em jogo é de monta, ou estão envolvidos avultados recursos, designadamente em intervenções financiadas, as entidades financiadoras poderão pressionar para a realização de estudos de avaliação em profundidade. Em segundo lugar, se uma organização recolhe regularmente informação sobre os seus resultados, isso melhorará as condições para a realização de estudos de avaliação, facilitando o papel das pessoas que fazem avaliação, que poderão aceder à informação sobre os resultados “pronta a usar”.

¹⁸ No relatório do grupo de trabalho de avaliação do Comité de Apoio ao Desenvolvimento da OCDE, relativo à revisão da experiência da gestão orientada para os resultados nas agências de cooperação para o desenvolvimento, é mencionada a preocupação no seio das agências doadoras de que as tarefas de medição e de reporte sobre o desempenho possam estar a “excluir” as avaliações, isto é, que ambas as abordagens possam estar a competir pelos cada vez mais escassos tempo e recursos.

argumentando que a medição de desempenho, além ser associada ao controlo por parte daqueles cujo desempenho está a ser medido, coloca, em geral, um **grande enfoque na eficiência, excluindo outros valores e medidas**. Estas circunstâncias, aliadas à visão da medição de desempenho como uma forma simplista e grosseira de produção de conhecimento (Greene, 1999), levam alguns membros da comunidade de avaliação a demitir a medição de desempenho dos processos de avaliação.

Apesar de, na comunidade de avaliação, a gestão e medição de desempenho serem encaradas com alguma relutância, encontramos argumentos importantes e persuasivos no sentido de considerar medição de desempenho e avaliação processos complementares. Foi na linha deste entendimento que Nielsen e Ejler (2008) propuseram várias ferramentas e processos de avaliação que podem ser adotados com vantagem no enquadramento da gestão de desempenho. Estas concetualizações sugerem formas de perceber se as organizações usam (ou não) dados de medição do desempenho e avaliação de forma coerente. É importante ressaltar, todavia, que a complementaridade é uma via de dois sentidos: as práticas de monitorização podem informar os estudos de avaliação e vice-versa. Como notam Nielsen e Hunter (2013: 120-121), o alcance de uma complementaridade desejável e útil entre avaliação e medição de desempenho convoca uma abordagem mais estratégica de produção de conhecimento e gestão tanto a nível organizacional como político. Isto implica também o reconhecimento entre quem toma decisões de que os sistemas de medição de desempenho e avaliação precisam efetivamente de ambos os elementos para se reforçarem reciprocamente.

Feito o esboço geral dos diversos tipos de abordagens à avaliação, assim como algumas das suas dimensões críticas e questões essenciais, no segundo capítulo situo a avaliação no contexto das organizações do terceiro setor. Além de fundamentar a relevância da avaliação nestas organizações e de traçar o panorama atual do setor em Portugal, percorrerei os principais modelos para explorar a sua eficácia e explorarei as relações entre avaliação, *accountability* e aprendizagem organizacional.

Capítulo 2. Avaliação nas organizações do terceiro setor: contextos e abordagens

2.1. Governação, desestatização e projetificação das políticas sociais

O debate sobre a crescente relevância da avaliação nas OTS inscreve-se, por um lado, nas discussões sobre (i) as **transformações no papel do terceiro setor** associadas às mudanças no papel e forma do Estado de bem-estar, como seja o seu crescente envolvimento nas políticas públicas, e, por outro lado, (ii) nas reflexões sobre as próprias **transformações nas políticas públicas**, como é o caso da nova gestão pública e da projetificação das políticas sociais.

(i) Transformações no papel do terceiro setor

Os estudos do terceiro setor¹⁹ procuram explicar a existência deste e demarcá-lo um espaço próprio que o distinga dos setores público e privado lucrativo. Desde a década de 1970 que, em termos internacionais, a referência ao terceiro setor está ligada às discussões sobre o Estado-Providência, ou seja, sobre o papel do Estado e da sociedade civil na integração social e no bem-estar das populações (Ferreira, 2013: 161). É no contexto da crítica e crise do Estado Providência que o terceiro setor emerge enquanto setor, descrito como incluindo organizações e práticas que não são do Estado nem do mercado e cujos equivalentes, localizados em contextos históricos específicos, como setor não lucrativo, setor voluntário, economia social e terceiro sistema, seguiram mais ou menos a mesma forma de autodescrição, como setor, e como diferente de outros setores (Ferreira, 2008).

Desde os anos 1980, a par da erosão do Estado-Providência, as OTS foram sendo encaradas como parceiros privilegiados de intervenção. As próprias linhas das orientações para as políticas provenientes por via da *soft law*²⁰ da União Europeia, bem como dos programas europeus, afirmam a importância das organizações do terceiro setor na

¹⁹ O conceito de terceiro setor – que será explicitado na secção seguinte - é aqui usado de modo equivalente a outros termos tais como setor voluntário, setor não lucrativo, economia social e organizações não-governamentais, embora cada um deles exprima aspetos específicos e radique em tradições históricas, sociais, políticas e culturais distintas.

²⁰ Conjunto de regras de conduta que, embora não possuam força legal, têm, ainda assim, efeitos práticos (Snyder, 1993 *apud* Valadas, 2012: 231). É o caso das recomendações e das resoluções da CE (em substituição das directivas e dos regulamentos, conhecidas por *hard law*).

governança em termos dos objetivos estratégicos de crescimento económico sustentado e coesão social. Este crescimento foi alimentado por uma crença entre as entidades financiadoras de que as OTS são mais eficientes e efetivas do que os governos na prestação de serviços sociais básicos, são mais competentes no apoio à população carenciada, e são atores-chave nos processos de democratização (Edwards e Hulme, 1996). O interesse crescente nas potencialidades do terceiro setor associa-se, assim, não apenas à necessidade de compensar as deficiências do Estado e do Mercado, mas também à necessidade de promover a participação dos cidadãos na construção de uma sociedade mais justa e democrática e de garantir níveis adequados e generalizados de bem-estar (Santos *et al*, 2015: 42).

Muita da discussão em torno do **conceito de governança**²¹ pretende apontar para **transformações na natureza do Estado** e, desta, muita pretende enfatizar a diminuição do papel direto do Estado, equacionado na forma específica do Estado-Providência ou de bem-estar (Ferreira, 2009). Os Estados e as suas políticas estão cada vez mais envolvidos em interações transnacionais, promovidas por agentes muito diversos, estatais e não estatais, que implicam um debate mais aberto e complexo. O que eram anteriormente os papéis indisputados do governo são agora vistos como “problemas sociais mais comuns e genéricos que podem ser resolvidos por instituições políticas mas também por outros atores” (Pierre, 2000: 4). Para muitos/as autores/as e discursos, **governança em rede** significa um reforço da oportunidade de participação das organizações do terceiro setor e dos cidadãos nas decisões políticas, na produção de serviços e na sua monitorização, em suma, na construção do “bem público” (Ferreira, 2012).

As novas tendências nos processos de governança têm representado uma oportunidade para a afirmação do terceiro setor enquanto ator político e económico.

²¹ A origem do conceito de governança é encontrada na ideia de orientação (*steering*), no sentido de guiar um navio, ou pilotar qualquer coisa, coisa essa que tanto pode ser uma organização, como um território, ou uma sociedade (Ferreira, 2012). No que se refere ao seu uso atual, e numa perspetiva mais restrita, o conceito (também designado de *governança*), refere-se a um modo específico de coordenação de atividades complexas e interdependentes, envolvendo redes de atores estatais e não estatais (Jessop, 2002), chamando a atenção para a participação de um número alargado de atores e interessados nos processos de *steering* (ou governança, no sentido lato). Pretende assinalar uma mudança nos modos de coordenação da centralidade das formas hierárquicas ou do mercado para a noção de redes, comunidade ou associação (Ferreira, 2012).

(Santos *et al.*, 2015). Com a chamada passagem do governo para a governação (Rhodes, 2000), o **terceiro setor veio a tornar-se parceiro do Estado na partilha de responsabilidades públicas não só pela provisão, mas também pela governação do bem-estar social** (Ferreira, 2012). Por um lado, as suas funções crescentes na criação de emprego e provisão de bens e serviços sociais destinados, por exemplo, ao desenvolvimento local, à diminuição da exclusão social ou à regeneração urbana. Um exemplo desta dinâmica é o das parcerias público-privadas envolvendo o Estado e o terceiro setor – que por vezes também incluem as empresas – com o fim de levar a cabo atividades de interesse público. Uma das formas comuns de parceria entre o governo e as OTS são os protocolos e acordos de cooperação no âmbito da segurança social, assinados entre o governo e as organizações federativas e de base, visando o fornecimento de um conjunto de serviços sociais. Por outro lado, nos debates sobre a governação ou governação em rede, o terceiro setor ocupa um lugar proeminente já não apenas como prestador de serviços em parceria com o Estado, mas também na organização dos serviços, tornando-se interlocutores e parceiros das políticas governamentais (Ferreira, 2012; Santos *et al.*, 2015).

As OTS são agora simultaneamente fornecedores de serviços, em grande medida subcontratados pelo Estado, e parceiros na governação com o Estado, representantes de públicos desfavorecidos, especialistas na provisão e necessidades, e parceiros do Estado na prossecução do interesse geral. Em Portugal esta tendência verificou-se sobretudo a partir de meados da década de 1990, através de políticas que envolveram o terceiro setor lado a lado com órgãos públicos e governamentais na conceção, desenvolvimento e implementação de políticas na área do emprego, luta contra a exclusão social, educação, proteção de crianças e jovens e desenvolvimento local, num modelo que ficou plasmado no Pacto de Cooperação para a Solidariedade Social, em 1996 (Ferreira, 2013; Santos *et al.*, 2015: 42). Os programas de intervenção cofinanciados por **fundos comunitários** (nomeadamente pelo Fundo Social Europeu), além de se revelarem essenciais no nosso Estado Providência fraco, podem ser identificados com estas novas tendências dos processos de governação, nomeadamente ao promover a ideia de governação multinível e

de parceria, preconizando o envolvimento de uma grande diversidade de atores - públicos, privados e do terceiro setor – na sua concepção e execução.

Estado, mercado e sociedade civil associam-se de forma diversa na resolução dos novos problemas e na criação de novas oportunidades (Almeida, 2011). Nestes novos instrumentos de governação, o Estado assume, por um lado, o papel de parceiro igual na governação e, por outro lado, o de metagovernador (Ferreira, 2011), ou seja, organizador das condições de governação (Jessop, 2003), pois a ele compete a **definição dos dispositivos reguladores**.

No contexto da amplificação de papéis e expectativas sobre o terceiro setor, aumenta a pluralidade de *stakeholders*, oriundos do próprio setor, do Estado, do mercado ou da comunidade, e constitui um desafio constante na medida em que destes provêm diferentes expectativas e exigências de prestação de contas (membros e associados, trabalhadores/as, voluntários/as, utilizadores/as, clientes, fornecedores, financiadores, governos, comunidades, federações e outras OTS, etc.).

A assunção do carácter público dos serviços sociais fornecidos pelas OTS, quer por efeito das próprias exigências das OTS relativamente a essa mesma responsabilização pública quer por transferência da provisão pública para o terceiro setor, à medida que estes serviços ocupam um lugar mais proeminente nas políticas sociais ou da contratualização do terceiro setor pelo Estado para o fornecimento destes serviços, traz **novas exigências de produção e de prestação de contas** (Ferreira, 2012: 108). Cada vez menos o Estado se concebe como fornecedor direto e mais como promotor e regulador, estendendo, por isso, essa regulação aos serviços fornecidos pelas OTS (*ibidem*). Nesta linha, a avaliação surge como elemento indispensável na contratualização entre Estado e terceiro setor, tendo em vista a implementação de programas sociais.

(ii) Transformações nas políticas públicas

- *A adoção dos princípios da nova gestão pública (NGP)*

A **orientação para a nova gestão pública no Estado**, carreando princípios de mercado para prestação de bens e serviços públicos, intensificou o controlo exercido sobre o próprio TS, estendendo à relação entre o Estado e as OTS critérios de qualidade, transparência e prestação de contas. As exigências de prestação de contas e medidas de

desempenho em instituições públicas, que caracterizam a NGP, acabam por ser transferidas para a relação do Estado com o terceiro setor, por via múltiplas formas de colaboração entre o Estado e as OTS.

Baseando-se o *ethos* da NPM nos princípios da economia, eficiência e eficácia, com uma ênfase específica na responsabilização pelos resultados, os agentes da administração pública são pressionados a demonstrar os resultados de suas ações e do trabalho dos setores/organismos que representam/lideram. Uma vez que em muitos domínios de atividade (bem-estar, saúde, educação, desenvolvimento local, etc.), os organismos públicos atuam através da intervenção do terceiro setor, esta nova filosofia de gestão implica exigências crescentes para que estas organizações demonstrem os resultados – e, mais recentemente, o impacto (social) – dos seus programas e serviços (Bassi, 2015). Behn descreve, exatamente, a NGP como "todo um conjunto de táticas e estratégias que procuram melhorar o desempenho do setor público - para melhorar a capacidade dos organismos governamentais e seus colaboradores dos setores sem e com fins lucrativos de produzir resultados" (2001: 26).

Sob influência da lógica da NGP vem sendo colocada uma maior ênfase não apenas na medição mas também na comparação do desempenho. Neste contexto, os mecanismos de **gestão da qualidade** (frequentemente emprestados do setor industrial) têm-se tornado uma ferramenta privilegiada para melhorar a produtividade nos serviços sociais (Huber *et al.*, 2008), e os respetivos padrões critérios a partir dos quais as propostas para a celebração de contratos públicos de prestação de serviços sociais são avaliadas e escolhidas (Mass e Rodrigues, 2010). A maior ênfase na qualidade dos serviços sociais encontra-se alicerçada no debate sobre os Serviços Sociais de Interesse Geral na UE e no argumento da maior eficiência na prestação de serviços e das maiores possibilidades de escolha para os utilizadores (*ibidem*). Tal relevo tem orientado as OTS para implementação de processos e normas de qualidade, capazes de as tornar competitivas em face das organizações lucrativas.

Algumas análises notam também que uma das implicações do movimento da nova gestão pública é a proliferação de métodos de gestão de projetos no setor público, uma

vez que aquele movimento conduziu ao recurso crescente a projetos, enquanto soluções organizacionais (Jałocha, 2013; Packendorff e Lindgren, 2014).

- *A projetificação das políticas sociais e os fluxos financeiros com a União Europeia*

Além de implicarem uma ação descentralizada e a partilha de responsabilidades e de ação com as organizações da sociedade civil, **os novos modelos de intervenção social emergentes do Estado-Providência** **carreiam uma orientação de abordagem “por projeto”**, a qual traduz uma mudança de perspectiva assumida hoje pelas políticas sociais. O projeto é a nova unidade básica de intervenção e constitui um requisito do princípio da gestão por objetivos. Enquanto a forma tradicional de intervenção se baseava nos planos de atividades elaborados pelos serviços públicos e correspondia a um modelo de regulação autónoma do Estado, os projetos integrados correspondem a um modelo de regulação partilhada entre o Estado e a sociedade, são suportados por fundos de proveniência pública e privada e submetidos a uma estrutura mais autónoma de responsabilidades (Hespanha, 2008: 4).

Quadro 5. As metodologias da intervenção social pública

	Abordagem por serviço	Abordagem por projeto
Unidade de gestão	O serviço público.	A parceria público-privado.
Duração	Permanente.	Por tempo determinado.
Administração	Rígida/burocrática.	Flexível/convencionada.
Base normativa	Lei orgânica.	Contrato de adesão a um programa.
Metodologia	Plano de atividades de acordo com o programa do governo.	Gestão por objetivos.

Fonte: Hespanha, Pedro (2008). Políticas Sociais: novas abordagens, novos desafios? *Revista Crítica de Ciências Sociais*, 39: 1, Quadro 1.

A importância de atividades temporárias, nomeadamente ao nível da intervenção social, é crescente, e a lógica de projeto encontra-se generalizada, não só na atuação dos organismos públicos e empresas, mas também na das organizações do terceiro setor. Esta tendência tem sido abordada por diversos autores, a diferentes níveis – societal, organizacional e individual (Kuura, 2011) – e sob diferentes designações (projetização, projetificação, orientação para projetos, etc.).

Na literatura acadêmica produzida na Europa, a **projetificação** – definida por Bredin e Söderlund (2011: 9) como a “transição da produção repetitiva para processos de trabalho não rotineiros e o recurso a projetos temporários” – **tem sido amplamente associada aos fluxos financeiros com a União Europeia** (Kovach e Kucherova, 2006; Kuura, 2011; Jałocha, 2013). Kovách e Kucherova (2006, 2009) referem que as transferências dos fundos comunitários, a par de fundos nacionais e outros, assentes na implementação de programas e projetos, reverteu numa “projetificação ‘bottom-up’” das políticas europeias e nacionais de desenvolvimento (regional). A projetificação da UE proporcionou, segundo os autores, a emergência de uma “classe de projeto” na Europa Central e Oriental, cuja função geral é de mediação na redistribuição dos fundos de desenvolvimento públicos e privados e de transferência de materiais, ideias, conhecimento e poder.

Algumas das análises focam, mais concretamente, as mudanças operadas no setor público em consequência da projetificação induzida pela política regional da UE e aplicação dos fundos estruturais europeus. Godenhjelm, Lundin e Sjöblom (2011, 2015), por exemplo, descrevem a forma como as políticas regionais da UE são aplicadas, desvelando os mecanismos pelos quais se tornam motores da projetificação no setor público europeu. Os autores concluem que sem mecanismos de interligação contextualmente sensíveis entre as estruturas públicas temporárias e permanentes, os projetos correm o risco de perder as suas qualidades flexíveis e inovadoras e de fragmentar a capacidade das organizações permanentes para manterem a coordenação e continuidade. Jałocha (2013), por seu turno, procurando compreender as consequências da projetificação nas organizações do setor público polaco, conclui que a projetificação – fortemente induzida pela UE – provocou mudanças de natureza diversa: desde a acomodação dos objetivos organizacionais aos objetivos dos pedidos de subvenção, às mudanças nas estruturas organizacionais. O estudo mostra que o facto de os projetos financiados pela UE serem concluídos dentro do prazo, não excederem o orçamento, e alcançarem os indicadores dos programas (seguindo o triângulo da gestão de projetos – âmbito, duração e custo), nem sempre é sinónimo de criação de valor para as comunidades locais, relevando a necessidade de a UE não avaliar o sucesso dos projetos apenas a partir de medidas de

tempo, custo e qualidade, mas também mediante o impacto local e valor criado para os *stakeholders*.

As consequências da projetificação, a nível organizacional (nomeadamente nas estruturas organizacionais e de governação), nos vários setores (público, empresarial e terceiro setor), têm merecido alguma reflexão na literatura. Packendorff e Lindgren (2014: 12) sistematizam algumas das consequências negativas da projetificação, quer a nível organizacional, quer individual: o potencial risco de re-burocratização; a desconsideração da necessidade de integração dos projetos em programas ou portfólios; o tempo limitado para o desenvolvimento de competências; o *stress* para o cumprimento de prazos; e ausência de confiança e de continuidade social. Algumas pesquisas recentes citadas pelos/as autores/as (*ibidem*: 14) também destacam as consequências emocionais do trabalho no âmbito dos projetos, retratando projetos como emocionalmente pesados e *espaços* potencialmente viciantes e nocivos. A partir de uma conceção *vasta* de projetificação – enquanto fenómeno cultural e discursivo – Packendorff e Lindgren (2014: 14) consideram também a forma como as aspirações e expectativas associadas aos projetos são tomadas como garantidas no trabalho dos projetos já que, através da definição de metas e do planeamento, as aspirações e expectativas são projetadas para o futuro, quase até ao ponto em que o futuro é "vivido" com antecedência, dado como certo e garantido. Ser alguém bem-sucedido numa sociedade projetificada é ser disponível, flexível e conectado, sacrificando projetos de vida, condições de estabilidade e de previsibilidade social (Chiapello e Fairclough, 2002, *apud* Packendorff e Lindgren (2014: 14).

Na sua análise sobre as tendências de projetificação, Kuura (2011: 133) conclui que embora o financiamento temporário (por projeto) não seja genericamente negativo, as atividades e organizações permanentes continuam a ser necessárias e devem ser suportadas por alocações de carácter permanente, sob pena de um grande número de pseudo-projetos serem continuamente usados para financiar atividades permanentes (ou de longo prazo) com fontes temporárias (ou de curto prazo), comprometendo a sustentabilidade das organizações.

Ainda que o fenómeno da projetificação seja difícil de estimar quantitativamente (não se reflete em estatísticas e tem sido pouco estudado em termos extensivos), várias

análises têm dado conta do **crescente peso do financiamento temporário** (baseado em projetos) nas organizações que beneficiam de **fundos públicos** (Kuura, 2011; Godjenheim, Sjöblom e Lundin, 2011; Jalocha, 2013). Isso tem preocupado diversos setores de atividade, incluindo o cultural, de desenvolvimento e o social. No nosso país, a evolução das despesas inscritas no domínio da Ação Social da Conta de Segurança Social parece sustentar esta tendência. Com efeito, entre 2008 e 2014, o peso relativo das despesas em *Programas e projetos* subiu de 4,8% em 2008, para 7,6% em 2014, constituindo a segunda rubrica mais representativa da despesa em ação social (liderada pelos *acordos de cooperação*, que representam perto de 80% do total da despesa) (IGFSS, 2008, 2014).

O terceiro setor, para o qual os fundos públicos (nacionais ou comunitários) constituem fontes importantes de financiamento, será particularmente vulnerável a esta tendência. Em Portugal, além das transferências e subsídios de caráter mais ou menos regular (com destaque para a atribuição de subsídios públicos fixada em função dos serviços prestados no âmbito dos acordos de cooperação), o financiamento público às OTS tem-se feito mediante a abertura de linhas de financiamento para projetos (ex. cooperação, educação e ação social), muitas das quais suportadas por fundos comunitários. Os **fundos estruturais europeus** são, pois, uma importante fonte de financiamento para muitas instituições e para algumas respostas na área social, por via da implementação de projetos²². Por exemplo, a nível nacional, o Quadro de Referência Estratégica Nacional (QREN)²³, que vigorou entre 2007 e 2013, permitiu o investimento em território nacional de mais de 9.500 M€ milhões de euros direcionados especificamente para a Agenda Temática Potencial Humano que, por se encontrar vocacionada para o desenvolvimento de projetos nos domínios da formação, cidadania, inclusão e desenvolvimento social (entre outras áreas), motivou o **envolvimento de um grande número de OTS na promoção de milhares de projetos.**

²² Aliás, o nascimento de muitas OTS (especialmente Associações de Desenvolvimento Local) foi despoletado pela criação de apoios comunitários, e dos fundos estruturais em particular. Como refere Quintão (2011), a estabilidade e aproximação aos padrões e às dinâmicas europeias do terceiro setor, conseguida a partir de 1986/1987 foi estimulada pelo mecanismo dos fundos estruturais, verificando-se um crescimento significativo do número de instituições.

²³ O Quadro de Referência Estratégico Nacional (QREN) constitui o enquadramento para a aplicação da política comunitária de coesão económica e social em Portugal no período 2007-2013.

As implicações do modelo de intervenção/financiamento por projeto nas organizações do terceiro setor têm sido aludidas em diversos estudos empíricos recentes. Por exemplo, no *Diagnóstico das ONG em Portugal*, coordenador por Raquel Franco (2015: 35), as ONG abrangidas pelo estudo identificam uma série de problemas associados ao financiamento com base em projetos candidatados a sistemas de incentivos públicos nacionais, ou europeus, designadamente o facto de “tornar o trabalho no terreno dependente de prioridades de agenda” (que podem não ser coincidentes com as suas), e de restringir a continuidade das intervenções, por se tratar de um financiamento “pontual e irregular”. No mesmo sentido apontam os resultados do estudo de diagnóstico do projeto “SPOTS -- Sustainable Projects for Organisations of the Third Setor” sobre o “Financiamento do terceiro setor em Portugal”, no âmbito do qual foram auscultadas entidades financiadoras do terceiro setor, entidades de capacitação e organizações beneficiárias. O estudo conclui, de forma contundente, que o “financiamento focalizado em projetos e não em organizações é limitativo da sustentabilidade das organizações, sendo o ciclo do projeto muitas vezes curto e desalinhado da estratégia geral da organização”, e cria “uma artificialidade em que a conveniência do processo para o financiador (...) se sobrepõe às reais necessidades da organização e dos seus beneficiários” (Ghanem, Pedra e Matos, 2016: 23).

Em todo o caso, **uma das etapas essenciais usualmente associada ao ciclo de um programa ou projetos é a avaliação** (e.g., Schiefer *et al.*, 2006; Capucha, 2008). Donde, o desenvolvimento de projetos ao abrigo de fundos públicos ou filantrópicos implica, em geral, a conformidade com uma série de requisitos de avaliação, os quais prescrevem lógicas, princípios e metodologias com capacidade para moldar as práticas de avaliação das organizações beneficiárias. No caso dos projetos ao abrigo de fundos estruturais europeus, de que se ocupa mais especificamente este trabalho, o potencial de influência dos requisitos de monitorização e avaliação nos modelos de avaliação exibidos pelas organizações é particularmente significativo, se atentarmos na **relevância das receitas de iniciativas sustentadas por fundos comunitários na estrutura de receitas das OTS** em Portugal (e.g., Ferreira *et al.*, 2007; Martín *et al.*, 2014).

2.2. O panorama atual do terceiro setor em Portugal

O termo terceiro setor, que perfilho neste trabalho, tem sido usado como fórmula genérica para descrever um conjunto de relações sociais diferentes das do Estado e do mercado. Ele aplica-se a um conjunto difuso de iniciativas, que podem assumir variadas formas e estatutos, mais ou menos formais, desde as cooperativas, associações, fundações, mutualidades, sindicatos ou organizações não-governamentais, até grupos de autoajuda, iniciativas populares de base, redes e movimentos sociais.

Alicerçadas em perspectivas e tradições teóricas distintas, têm sido apresentadas listas de atributos para definir e classificar o terceiro setor. Estas podem referir-se a características organizacionais – como o carácter privado, a ausência de finalidades lucrativas, o objetivo de beneficiar a comunidade ou os seus membros, o seu carácter autogovernado e voluntário e o seu grau de organização formal – ou a racionalidades ou valores específicos – solidariedade, participação democrática, enraizamento local (Ferreira, 2009: 322).

Terceiro setor é apenas uma denominação recente, dentre muitas designações, que se inscrevem não só em diferentes tradições culturais e históricas dos centros de produção de conhecimento, mas também em diferenças efetivas das próprias realidades sociais estudadas nos diversos países e territórios²⁴. Embora se trate de um conceito contestado, designadamente por ser demasiado abrangente ou por se prestar a equívocos ao tender a confundir-se com o setor terciário (Parente e Quintão, 2014), o seu carácter abrangente e flexível permite dar conta de realidades heterogéneas e complexas, o justifica a opção desta pesquisa em assumir o terceiro setor como campo concetual. De resto, visando abarcar realidades diferenciadas, o conceito de terceiro setor tem sido particularmente útil na investigação comparativa internacional, procurando ultrapassar os limites das tradições teóricas mais instituídas e integrar a diversidade das realidades nacionais e regionais (Quintão, 2011: 6-7). Daí que seja usado como um equivalente para definições mais locais e contextuais, como setor voluntário (Reino Unido e países nórdicos), setor não lucrativo (EUA), economia social (Europa continental), ONG e organizações de

²⁴ Sobre as diversas abordagens teóricas e conceptuais sobre o terceiro setor, consultar Chaves e Monzón (2007) e Evers e Laville (2004).

base (países do Sul), terceiro sistema (União Europeia), etc. (Ferreira, 2009: 322). Acompanhando o entendimento de Sílvia Ferreira (*ibidem*: 322), independentemente da designação ou do seu conteúdo específico, o que é relevante “é o surgimento de um setor em relação com o Estado e o mercado, como um “outro” distinto da economia de mercado e do Estado, e a tentativa de observar aspetos comuns num amplo leque de iniciativas frequentemente isoladas entre si”. Segundo a mesma autora, mais do que outros termos alternativos, o de terceiro setor está intimamente ligado à discussão sobre o futuro do Estado-Providência e das políticas públicas.

Na linha do proposto por Ferreira (2009, 2015), a **definição de terceiro setor** é percebida, neste trabalho, como **contextual e relacional**. Contextual, uma vez que a identidade do TS varia no tempo e no espaço, em associação com mudanças ocorridas no seu ambiente, sendo a identidade das próprias organizações marcada pelos contextos da sua emergência (Santos et al., 2015: 36). Relacional, porque remete para a ideia de que ocupa uma área intermédia entre três setores: o Estado, o mercado e a comunidade/sociedade (Evers e Laville, 2004). Tal posicionamento situa o TS num campo de tensão - o espaço público - em que as organizações são simultaneamente influenciadas por políticas estatais e legislação, pelos valores e práticas das empresas privadas, pela cultura da sociedade civil, da família e da vida comunitária (*ibidem*: 15). Neste espaço, onde diferentes racionalidades e discursos confluem, as organizações posicionam-se distintamente em relação a cada um dos três pilares, recolhendo destas influências diversas, e podendo posicionar-se de forma diferenciada em diferentes momentos (Evers e Laville, 2004; Santos *et al.*, 2015). A noção das organizações do terceiro setor como **híbridas** (e.g., Evers and Laville 2004; Brandsen *et al.*, 2005), isto é, que misturam características em geral presentes separadamente nos diferentes setores, constitui a sequência lógica da conceção do TS como campo intermediário sob influência de princípios típicos dos sistemas do Estado, do mercado e da comunidade. Enquanto polivalentes e híbridas, as OTS misturam racionalidades, tarefas, e papéis diferentes (por exemplo, podem ser fornecedoras de serviços e canais de interesses e de pressão política), originando diferentes tipos de lógica e cultura organizacional (Santos *et al.*, 2015), mas também criando tensões nas relações

com os diferentes setores (Laville, 2005), as quais encontram expressão no conceito de isomorfismo institucional²⁵, o qual será retomado mais à frente.

Em todo o caso, enquanto parte de um espaço híbrido, e na ligação aos diversos sistemas, as OTS desempenham uma **multiplicidade de papéis** que traduzem o contributo do setor para a sociedade e **exprimem o seu impacto social** (Santos *et al.*, 2015). Dentre uma variedade de sistematizações sobre os **papéis do TS**, destaco a proposta de Neumayr e Meyer (2010) que, assumindo o carácter multifuncional do terceiro setor, apresentam um triângulo no qual associam as suas principais funções/papéis – prestação de serviços, advocacia e construção da comunidade – aos subsistemas económico, político e social:

- Na ligação ao **sistema económico** os autores apontam sobretudo os papéis de **prestação de serviços** (de ação social, saúde, educação, desenvolvimento local, etc.)
- Na ligação ao **sistema político** os autores identificam os papéis de **advocacia** (defesa de direitos ou de interesses de grupos específicos e esforços que visam a mudança de políticas ou a alteração da opinião pública);
- Na ligação à **comunidade** os autores identificam o papel de **construção de comunidade** (estabelecimento de laços e relações de proximidade e de confiança).
- Os autores identificam ainda um quarto papel, o de **inovação social** que, por ser **transversal** aos diferentes sistemas e papéis, não é integrado no triângulo.

Neste âmbito, retomo também a tipologia usada no âmbito do “The Johns Hopkins Comparative Nonprofit Sector Project (CNP)” (e.g., Salamon *et al.*, 2012) que, agrupando as organizações a partir das atividades que desenvolvem, distinguem dois tipos de funções do setor: **funções de serviço** e **funções de expressão**. As primeiras envolvem o fornecimento de serviços diretos como a educação, saúde, habitação, desenvolvimento

²⁵ DiMaggio e Powell (1991: 66) definem isomorfismo como “conjunto de restrições que forcem uma organização em determinado campo organizacional a assemelhar-se a outras organizações que enfrentam condições ambientais semelhantes”. Através do conceito de isomorfismo institucional, os autores descrevem o processo através do qual as OTS tendem a assumir características das organizações públicas na sequência da crescente assunção do carácter público dos serviços sociais fornecidos pelas OTS ou da contratualização pelo Estado para o fornecimento destes serviços, gerando novas exigências de produção e de prestação de contas (Ferreira, 2012: 108).

local, entre outros; enquanto as segundas envolvem atividades de expressão de valores, interesses e crenças culturais, espirituais, profissionais ou políticas, abrangendo instituições culturais, grupos de recreação/lazer, associações profissionais, grupos de defesa de causas, organizações comunitárias, grupos de direitos humanos, movimentos sociais, entre outros.

Nas últimas décadas, o crescimento, multiplicação e diversificação dos papéis e intervenções das organizações do terceiro setor, o desenvolvimento da investigação científica sobre o tema e a crescente atenção por parte das políticas públicas (CESE, 2007: 22-26) tem-se refletido numa tendência para a visibilização histórica e institucional do terceiro setor.

Em Portugal, até recentemente, o panorama do **TS era altamente fragmentado**, com diferentes tipos de organizações e tradições (estruturados em plataformas, federações e redes), com alianças políticas, identidades e discursos normativos distintos, e com relações diferentes e desiguais com o Estado, resistindo ao reconhecimento da existência de outros campos do terceiro setor e ainda mais à ideia de terem algo em comum (Ferreira, 2015: 218). À exceção da Constituição da República – que afirma a existência de um setor social e cooperativo, a par de um setor público e de um setor privado – não tinham ainda sido criadas estruturas legais ou políticas amplas que acolhessem uma ideia geral de terceiro setor, a nível político. Do mesmo modo, não existiam organismos de tutela ou alianças abrangentes que promovessem esta ideia entre os intervenientes do setor (*ibidem*: 213).

Foi apenas no início desta década, em contexto de crise internacional, que um "setor" global despertou a atenção política (*ibidem*: 214). Inicialmente, com a aprovação do primeiro Programa de Apoio ao Desenvolvimento da Economia Social (PADES)²⁶, que veio consagrar um conjunto articulado de medidas de estímulo ao desenvolvimento da economia social, entre as quais a criação de novos organismos de tutela da economia social e depois com a primeira lei de bases da economia social (Lei 30/2013, de 8 de Maio). **O terceiro setor português encontra-se agora estruturado sob uma versão abrangente do conceito de “economia social”**. De acordo com a lei de bases, integram a economia social,

²⁶ Resolução do Conselho de Ministros n.º 16/2010, de 4 de Março.

cooperativas, associações mutualistas, misericórdias, fundações, instituições particulares de solidariedade social (IPSS), associações culturais, recreativas, desportivas e do desenvolvimento local, além das entidades abrangidas pelos subsetores comunitário e autogestionário, integrados nos termos da Constituição no setor cooperativo e social e outras entidades que respeitem os princípios orientadores da economia social e constem da base de dados da economia social. A Conta Satélite da Economia Social (CSES) alarga ainda mais o campo da economia social incluindo organizações que não estão recobertas pela própria definição da Lei de Bases, como entidades religiosas ou associações sindicais, patronais e políticas (Ferreira, 2015). O conceito enquadra não apenas as definições dominantes do “modelo” da economia social (corrente francófona), como sejam mutualidades, cooperativas e associações²⁷, mas também entidades dominantes do “modelo” não lucrativo (corrente anglo-saxónica), como é o caso das IPSS, definidas pelo estatuto jurídico especial que lhes é atribuído por prosseguirem um conjunto de atividades enquadradas na segurança social. Além disso inclui todas as cooperativas e mutualidades, ao contrário do modelo anglo-saxónico, e as organizações religiosas, sindicais e outras que no modelo francófono não estão habitualmente incluídas.

O presente trabalho acolhe a diversidade de entidades consideradas pela lei-quadro portuguesa, entendendo o terceiro setor como sendo diferente do setor público e do setor privado lucrativo, integrando organizações com práticas heterogéneas e difusas, orientadas para finalidades de bem-estar coletivo ou público, sem a primazia de uma finalidade lucrativa. Embora acolhendo a amplitude do campo coberta pela CSES, a tipologia de entidades adotada nesta investigação para discussão dos seus resultados é distinta, no sentido de melhor captar as formas identitárias do terceiro setor. Assim, tendo presente as características do terceiro setor português, as organizações foram agrupadas a partir do cruzamento de informação sobre a sua forma jurídica, estatuto e vocação principal:

²⁷ As mutualidades e cooperativas (excluídas do conceito de setor não lucrativo) permitem, até determinado ponto, o retorno de excedentes económicos não reinvestidos na missão.

- ADL (Associações de Desenvolvimento Local) – organizações vocacionadas para a promoção e a rentabilização de recursos em determinadas áreas geográficas, numa perspetiva integradora.
- Coletividades – Coletividades e outras associações de cariz popular que desenvolvem atividades nas áreas da cultura, recreio e desporto.
- IPSS (Instituições Particulares de Solidariedade Social) – Organizações com estatuto de Instituições particulares de solidariedade social (de natureza associativa, fundacional ou cooperativa).
- ONG (Organizações Não-Governamentais) – ONGD, ONGA, ONGPD, ONGM e outras organizações cívicas e de defesa de causas (associações empresariais, sindicais, políticas, etc.).
- Outras – entidades de natureza e estatutos diversos não abrangidas pelas categorias anteriores (escolas profissionais e outros estabelecimentos de ensino particular e cooperativo; cooperativas de comercialização, construção, serviços, etc.; casas museu, companhias de teatro e outras entidades culturais; bombeiros, ...).

Os diversos exercícios da quantificação do terceiro setor em Portugal (Carvalho, 2010; Salamon *et al.*, 2012; INE e CASES, 2013, 2016) revelam que, em termos gerais, o terceiro setor tem ganho uma **importância crescente na economia**, quer ao nível da **geração de emprego** e de fornecimento de produtos e serviços, quer em termos de **receita**, promovendo a dinamização social e económica do país. A análise de Ana Carvalho (2010) mostra que no espaço temporal de uma década, entre 1997 e 2007, o número total de organizações cresceu cerca de 90% e o nível de emprego no setor aumentou quase 85%. De acordo com a segunda edição da Conta Satélite da Economia Social (CSES)²⁸, no ano de 2013, o terceiro setor português engloba um total de 61.268 organizações, que

²⁸ Publicada pela primeira vez em 2013, a Conta Satélite da Economia Social foi elaborada no âmbito do Protocolo de cooperação assinado em 2011 entre o Instituto Nacional de Estatística (INE) e a Cooperativa António Sérgio para a Economia Social (CASES), com o objetivo de avaliar de forma exaustiva a dimensão económica e as principais características da Economia Social em Portugal. A CSES, para 2010, constituiu a última etapa no processo de desenvolvimento de uma conta alargada e completa para a Economia Social, que se iniciou em 2011 com a apresentação da Conta Satélite das Instituições sem fim lucrativo (CSISFL) para 2006, na denominada “versão reduzida” definida pelo *Handbook on Non-Profit Institutions in the System of National Accounts*. (INE e CASES, 2013).

empregaram aproximadamente 215.963 pessoas, o que representa 6% do emprego total remunerado a nível nacional (INE e CASES, 2016) e constitui um aumento relativamente ao ano de 2010. As associações de cultura, desporto e recreação (sobretudo associativismo de base popular) representam cerca de metade das organizações, seguidas da ação social e dos cultos e congregações, que representam, respetivamente, e 15,6% e 13,7% das iniciativas.

A análise do Valor Acrescentado Bruto (VAB)²⁹ e da distribuição do emprego remunerado por atividade revela uma hierarquização diferente da anterior, pondo a nu as **grandes assimetrias dentro do setor e a forte associação da economia social às funções do estado de bem-estar**. Com efeito, embora representem uma parte significativa do total das organizações do terceiro setor, o peso das organizações que exercem atividades na área da cultura, desporto e recreio em termos do Valor Acrescentado Bruto (VAB) e emprego remunerado (ETC) é relativamente diminuto (4,9% e 6,3%, respetivamente) (INE e CASES, 2016). Já as organizações que exercem atividade em áreas tradicionais de intervenção do estado social, que incluem a proteção social (ação social e saúde), habitação e desenvolvimento e educação, representam cerca de 60% do VAB total apurado para o TS e 64,6% das remunerações totais geradas, ainda que correspondam a apenas cerca de um quarto do total das organizações (*ibidem*). O peso significativo das Instituições Particulares de Solidariedade Social (IPSS) na composição global do setor – no universo de mais de 55 mil organizações existem mais de 5 mil com estatuto de IPSS – confirma a importância da proteção social assegurada por estas entidades, a partir da contratualização dos seus serviços pelo Estado.

Uma leitura dos dados disponíveis sobre as fontes de receitas das organizações do terceiro setor é indicadora da **prevalência dos recursos gerados pela produção** (vendas e serviços) na estrutura de receitas das organizações, ainda que se verifique uma diminuição do peso deste tipo de receita entre 2010 e 2013 (60,1% face a 62,8%). Esta diminuição é compensada pelo crescimento do peso das transferências e subsídios do Estado naquele período (26,7% face a 23,8%), embora este tipo de recurso continue a ter um peso bem

²⁹ Constitui o resultado líquido da produção avaliada a preços de base e diminuída do consumo intermédio avaliada a preços de aquisição. O valor acrescentado é registado a preços de base.

menos significativo no total das receitas geradas (INE e CASES, 2013, 2016). Não deixa de ser assinalável verificar que o peso dos recursos gerados pela produção é muito superior ao dos subsídios no total de receitas das organizações ligadas aos serviços sociais, saúde e educação (os subsídios representam, respetivamente, 31%, 25,4% e 32,3% das receitas). Sobretudo tendo em conta que estas áreas, estando às responsabilidades do Estado de bem-estar, são tipicamente associadas a níveis mais elevados de financiamento público, sobretudo em contextos em que as OTS assumem um papel central, como é o caso dos países da Europa continental e o de Portugal, em particular ao nível dos serviços sociais e da educação pré-escolar (Ferreira, 2015: 233).

Todavia, como é defendido em diversos trabalhos (Carvalho, 2010; Lopes, Parente e Marcos, 2014; Ferreira, 2015), **pode ser equivocado assumir** que o peso relativamente reduzido do financiamento do Estado nas receitas das OTS – em particular daquelas que operam ao nível do bem-estar (como sejam as IPSS) – indique **autonomia financeira perante o Estado**. Lopes, Parente e Marcos (2014) ressaltam que é necessário ser prudente na interpretação daqueles dados, considerando que apenas os setores fundacional, mutualista e cooperativo (fundamentalmente devido às organizações que se integram na área financeira), observam capacidade líquida de financiamento. Segundo dados da Conta Satélite da Economia Social (CSES), a necessidade líquida de financiamento é sobretudo afeta às associações e outras organizações da economia social e às misericórdias (sendo que parte significativa destas entidades tem estatuto de IPSS). De resto, as organizações com este estatuto jurídico representam 38,2% da necessidade líquida de financiamento da Economia Social. Razões que levam Lopes, Parente e Marcos (2014) a correlacionar estas formas jurídicas com a sua ausência de autonomia financeira. Por outro lado, como ressaltam estas autoras, o peso da produção enquanto recurso preponderante gerado por estas entidades “não deixa de significar uma dependência perante o financiamento estatal”, ainda que “sob a força de prestação de serviços” (*ibidem*: 94). Nesta linha, Sílvia Ferreira (2015: 224) prefere falar de uma **relação de “dependência mútua”**, tendo em conta que em Portugal apenas 2,5% dos serviços sociais são geridos pelo Estado e o elevado peso dos serviços de apoio à família, o que atesta a posição central que estas organizações ocupam no sistema de bem-estar.

Uma das razões por que as organizações de solidariedade social foram fortemente afetadas pela **crise económica** prende-se exatamente com a sua estrutura de financiamento, conforme avançado noutras análises (Ferreira, 2015; Franco, 2015; Santos *et al.*, 2015), e reiterado em entrevistas com representantes de organizações e das respetivas estruturas representativas no âmbito deste trabalho. Em face da **degradação da sua base de receitas**, decorrente do decréscimo nos rendimentos e do aumento do desemprego entre membros, utilizadores/as e beneficiários/as das OTS, assim como das medidas de austeridade (Ferreira, 2015), algumas análises apontam para a maior orientação das estratégias do setor para **um modelo de geração de receitas alinhado com as lógicas do empreendedorismo social** (Parente, Lopes e Marcos, 2012; Parente *et al.*, 2014), com o intuito de ganharem sustentabilidade através de práticas de gestão e recursos do mercado. Surgiram também alguns **novos fluxos de financiamento fora do quadro da relação com o estado de bem-estar**, tais como fundações filantrópicas, que têm vindo a financiar projetos inovadores de curta duração para resolver problemas sociais (Ferreira, 2015: 224). Os resultados do estudo de Santos *et al.* (2015) sobre o impacto da crise nas OTS vêm, todavia, moderar a amplitude dos efeitos destas novas tendências e atores na condição atual do setor. A análise revelou que, perante diminuição significativa do rendimento global no período da crise, nenhuma das diferentes fontes de receita teve capacidade de compensar a diminuição nas outras fontes de receitas, não se tendo verificado, por exemplo, uma substituição de receitas de fundos públicos por receitas de filantropia ou por receitas de vendas de bens e serviços. Isto leva as autoras a concluir que, “ao contrário do discurso de que as OTS podem encontrar “autonomia” relativamente ao financiamento público em atividades mercantis, não se afigura que o mercado ou a filantropia possam constituir uma alternativa viável para a sustentabilidade do sector” (*ibidem*: 229)³⁰.

Em todo o caso, alguns desenvolvimentos recentes ao nível da institucionalização do setor em Portugal, e das orientações de política europeia e nacional, apontam para um

³⁰ Exemplos destas iniciativas incluem o Programa EDP Solidária (<http://www.fundacaoedp.pt/inovacao-social/programa-edp-solidaria/programa-edp-solidaria/444>), o Programa Impacto social (<http://impactosocial.pt/>) e o Programa de Apoio a Empresas Sociais (<http://bancodeinovacaosocial.pt/empreendedorismo.php>).

reforço da ênfase na dimensão económica do terceiro setor. Isso é visível, por exemplo, na adoção do termo “economia social” por parte das novas instituições³¹ e na ênfase colocada na autossustentabilidade e na promoção do contributo do setor para a criação de emprego e para o crescimento, não obstante a imposição de limites à penetração das lógicas de mercado no TS (Ferreira, 2015). Ao nível das políticas, tem sido colocada a ênfase na **inovação e no empreendedorismo social**, percebidos sobretudo como a mimetização das práticas e estratégias do campo das empresas lucrativas e, mais recentemente, na **lógica do investimento social, por oposição às lógicas tradicionais de financiamento do terceiro setor** (donativos e subsídios). O acordo de Parceria Portugal 2020, estabelecido a 25 de julho de 2014, entre o Estado Português e a Comissão Europeia, identifica exatamente a inovação e empreendedorismo social como uma área de aposta fundamental para os fundos europeus estruturais e de investimento (FEEI), em particular para o Fundo Social Europeu. O governo português concretiza esta aposta da política europeia e nacional através da criação da iniciativa “**Portugal Inovação Social**” (Resolução do Conselho de Ministros n.º 73-A/2014), com o objetivo de contribuir para o desenvolvimento e promoção de um mercado de investimento social em Portugal, potenciado por instrumentos de financiamento inovadores e distantes da lógica tradicional de donativo, tais como o Fundo para a Inovação Social³² e os Títulos de impacto social³³.

O investimento social aparece associado ao **investimento pelo impacto**, uma vez que procura conciliar a geração de retorno financeiro e impacto social, sendo que um dos fatores fundamentais para informar a tomada de decisão em projetos de investimento social é exatamente a medição de impacto. Este interesse é ainda sinalizado pelo

³¹ Além do Programa de Apoio ao Desenvolvimento da Economia Social (PADES) e da lei de bases da economia social, em 2010 foram criadas mais duas estruturas de abrangência global do setor: a Cooperativa António Sérgio para a Economia Social (CASES), que sucede ao Instituto António Sérgio para o Setor Cooperativo (INSCOOP), e assume o papel de interlocutor público junto do terceiro setor e o Conselho Nacional para a Economia Social (CNES).

³² Financiamento de natureza grossista com fundos participados, para apoio a iniciativas e investimentos em inovação e empreendedorismo social em fase de consolidação ou disseminação, através da concessão de empréstimos, bonificação de juros, prestação de garantias ou quase-capital (Resolução do Conselho de Ministros n.º 73-A/2014).

³³ Apoios reembolsáveis contratualizados em parceria, para financiamento de soluções inovadoras na prestação de serviços públicos, orientadas para a obtenção de resultados e redução de custos (Resolução do Conselho de Ministros n.º 73-A/2014).

surgimento de iniciativas diversas em torno o impacto social, as quais têm contado com uma **forte representação de atores empresariais e fundacionais**. É o caso do Grupo de Reflexão para a Avaliação do Impacto Social (GRAIS)³⁴ e do Grupo de Trabalho Português para o Investimento Social³⁵, os quais já originaram documentos com linhas de orientação para entidades investidoras/financiadoras e recomendações para o crescimento do mercado de investimento pelo impacto em Portugal.

A orientação do financiamento para os resultados/impactos constitui, de resto, uma orientação transversal do novo ciclo de programação dos fundos europeus 2014 - 2020, que marca a deslocação do enfoque da operacionalização e implementação para os resultados e impactos, nomeadamente mediante a contratualização de resultados – consagração de financiamentos proporcionais à superação dos resultados contratados, mas também a penalização por incumprimento, total ou parcial de resultados³⁶.

2.3. Desafios da gestão das organizações do terceiro setor: dilemas da profissionalização

As crescentes pressões que o terceiro setor enfrenta no sentido de se tornar mais eficiente e "empresarial", com a ascensão da filantropia de "elevado envolvimento" (*high-engagement philanthropy*) e do empreendedorismo social, bem como de novas exigências de prestação de contas e de racionalização, tornam particularmente relevante a compreensão do fenómeno da profissionalização do setor (Hwang e Powell, 2009).

Genericamente, a profissionalização descreve a convicção de que as organizações deverão ser comandadas por especialistas (Salamon, 1999). Refere-se ao desenvolvimento de capacidades, identidades, normas e valores associados à pertença a um grupo profissional. Ao nível individual, é tradicionalmente medida pelo grau de escolaridade

³⁴ O grupo é composto por membros das seguintes entidades: AESE Escola de Direção e Negócios, CASES (Cooperativa António Sérgio para a Economia Social), Comunidade Vida e Paz, EVERIS, Fundação Calouste Gulbenkian, Fundação EDP, Fundação Montepio, GRACE, IPAV (Instituto Padre António Vieira), Instituto da Segurança Social, Santa Casa da Misericórdia de Lisboa, União das IPSS de Lisboa e a Faculdade de Ciências Humanas da Universidade Católica de Lisboa.

³⁵ Promovido pela Comissão Europeia, Fundação Calouste Gulbenkian, Laboratório de Investimento Social e *Social Finance*.

³⁶ Decreto-Lei n.º 137/2014 de 12 de setembro.

alcançado. No entanto, também pode ser medida a partir de padrões estabelecidos por uma associação profissional num determinado setor de atividade (Anheier, 2005: 205-206).

Nesta linha, a adesão à profissionalização poderá implicar desde o investimento no aumento dos níveis de qualificação do voluntariado, à contratação de pessoal remunerado e à maior ênfase nas habitações escolares (Lundstrom, 2001). Hwang e Powell (2009) distinguem *profissionalização gestonária*, que envolve mais pessoal com formação de base em gestão de negócios, e *profissionalismo em domínios substantivos*, tais como medicina, educação ou serviço social, o que, em alguns casos, pode ser antitético às práticas do negócio.

Os debates em torno da profissionalização da gestão das OTS têm sido alimentados por um contexto marcado por **desafios de ordem prática e teórica**. Prática, porque revelam um momento de mudança na configuração do setor e no modo como as organizações que o compõem têm sido geridas. Teórica, porque se inserem num domínio de produção de conhecimento em expansão, marcado pela proliferação de teorias, conceitos e modelos que procuram refletir as especificidades destas organizações, particularmente em torno das questões da prestação de contas, da governação e da gestão de pessoal remunerado e voluntário.

Os novos desafios que se colocam às OTS estão relacionados com mudanças políticas, económicas e sociais de nível internacional e nacional, que ocorreram sobretudo nas últimas três décadas, em que se presenciou um crescimento significativo do setor, tanto em dimensão e alcance das ações, como em popularidade (Herman *et al.*, 2005; Worth, 2014). Estas mudanças carregam preocupações de capacitação técnica e financeira destas organizações e de investimento nos seus sistemas internos (por exemplo, recursos humanos, processos e infraestruturas) e nas relações externas (por exemplo, financiadores, parceiros, voluntariado) no sentido de prosseguirem a sua missão e ampliar os impactos (Hudson, 2003: 37). Em ambientes incertos, as OTS sentem maior pressão para se profissionalizarem e para introduzirem maior controlo tecnocrático. Esta é também uma tendência em organizações que evoluem de uma condição de *start-ups*, de base voluntária, para entidades mais maduras, que podem ter que provar a sua legitimidade perante entidades financiadoras.

Se há três ou quatro décadas atrás as OTS eram quase sempre dirigidas por gente “amadora” – pessoas que acreditam numa causa e querem fazer algo por ela, mas que nem sempre têm as competências de gestão necessárias –, atualmente, estas organizações são geridas por profissionais, com o suporte de um corpo cada vez maior de conhecimento sobre as suas especificidades, por distinção em relação aos setores empresarial e público. Associadas ao trabalho com múltiplos *stakeholders*, as técnicas de gerir os diversos fluxos de financiamento para uma crescente diversidade de atividades, e de responder a um escrutínio público cada vez maior, são agora bem compreendidas por gestores e dirigentes experientes de muitas OTS (Hudson, 2003: 2).

Há **crecentes exigências** que as organizações do terceiro setor devem transpor para os resultados que alcançam, a que crescem as suas atividades e finanças, como sejam:

- a pressão para descobrir quais são os serviços que realmente fazem a diferença, para focalizar e expandir (“*scale up*”) as suas atividades no sentido de alcançar maior impacto;
- a pressão para formar alianças estratégicas com organizações congêneres e dos outros setores, para alcançar objetivos cada vez mais ambiciosos;
- a expectativa de que estas organizações se tornem mais sustentáveis;
- a maior regulação, por vezes de órgãos com requisitos diversos e conflitantes.

Por outro lado, as organizações do terceiro setor caracterizam-se por **padrões culturais próprios**, diferenciados dos que caracterizam as empresas e as instituições públicas, e que influenciam a forma como são planeadas e modeladas as práticas de gestão. Como sugerem Edwards e Folwer (2002), as organizações do terceiro setor são (ou assumem-se como sendo) sustentadas por **valores** (religiosos, políticos, humanitários, morais), o que complexifica a natureza dos objetivos, incentivos e estruturas de gestão, e reivindica uma perspetiva normativa da gestão que traduza a importância da adoção de valores na tomada de decisão quotidiana (Anheier, 2005). As OTS também atuam em **ambientes complexos e incertos**, exigindo abordagens que reconheçam e respondam às **necessidades de gerir múltiplas relações**, alternando combinações de papéis e fontes de recursos. A pluralidade das fontes de recursos torna também a gestão particularmente

complexa, dada a diversidade de regras e lógicas que encerram os diferentes recursos estatais, mercantis ou da comunidade Santos *et al.* (2015: 43). Além disso, como salienta Hudson (2003), estas organizações diferem das dos outros dois setores no sentido em que não existe uma ligação clara entre quem financiam e quem utiliza os serviços. No setor de mercado, as pessoas pagam pelos bens ao preço de mercado; se a organização não os fornece com qualidade e preço adequados, ou melhora o seu desempenho ou sai do mercado. No setor público, num sistema político-democrático, se as pessoas não acedem a serviços com um nível de qualidade aceitável podem, pelo menos em teoria, votar para retirar os agentes dos seus cargos. Para as organizações do terceiro setor, essa falta de responsabilização clara através dos mercados ou do processo político cria desafios e problemas de gestão complexos.

Estes **elementos distintivos do terceiro setor originam desafios distintos**, como sejam as dificuldades de monitorizar o desempenho, de gerir prestações de contas múltiplas, de intrincar as estruturas de gestão no sentido de equilibrar múltiplos constituintes, a que acrescem os conflitos entre voluntarismo e profissionalismo, a necessidade de não perder de vista os valores em que se fundam, e a tendência das OTS de estabelecerem objetivos vagos. Como é que uma organização se pode profissionalizar sem se tornar elitista ou exclusivista e sem perder de vista as pessoas beneficiárias? Como pode uma organização exigir profissionalização do pessoal voluntário? Será que a profissionalização do trabalho social voluntário envolve o risco de minar as dimensões ética e moral do voluntarismo?

Em face destes desafios, muitas OTS manifestam alguma **relutância em levar a sério os processos de gestão**, por razões diversas.

1. Um delas tem que ver com o facto de muitas organizações se caracterizarem por uma “cultura de ação”, na qual a liderança e o pessoal técnico resistem a ocupar um volume de tempo significativo a refletir sobre questões organizacionais, uma vez que essa priorização poderá interferir com a tarefa primordial de “trabalhar no terreno”.
2. Uma segunda razão é a visão partilhada por alguns segmentos sociais de que as OTS devem aplicar a quase totalidade dos seus fundos no trabalho de terreno e não

devem gastar tempo e dinheiro em aspetos administrativos. Como Smillie (1995: 151 *apud* Lewis, 2001: 9) sugere, existe um "mito poderoso de que o desenvolvimento deve ser barato".

3. Uma outra razão está ancorada no rápido crescimento e mudança que as OTS têm experimentado, que faz com que estejam sempre "um passo atrás" no pensar e agir em torno das respostas organizacionais (Lewis, 2001: 9). Com efeito, as OTS, que começam por constituir estruturas pequenas e informais, nas quais a gestão pode ser feita numa base *ad hoc*, informal, quando crescem rapidamente não se apercebem imediatamente de que precisam de novas ideias, sistemas e procedimentos em que se apoiar.
4. Uma quarta razão remete para forças e pressões externas. Uma vez que muitas organizações cresceram muito próximas das entidades financiadoras, foi-lhes sendo exigido que desenvolvessem sistemas de prestação de contas, podendo a sua eficiência e eficácia ser questionadas e desafiadas. Isto levou à percepção nas organizações de que muito do impulso para pensarem sobre a sua gestão vem de fora, e por isso é suspeito. As agendas de gestão tomaram por vezes a forma de gestão "imposta", ao invés de emergirem da agenda própria destas organizações (*ibidem*: 9).
5. Uma quinta razão que pode fundamentar a relutância com que a gestão tem sido associada às OTS baseia-se no receio do que Chambers (1994) designa de "profissionalismo normal", que se refere "ao pensamento, valores, métodos e comportamentos dominantes numa profissão ou disciplina" e que, "tal como a ciência normal, é conservador", baseando-se "numa estrutura de conhecimento e de geração de conhecimento centro-periferia, na educação e formação, na hierarquia das organizações e nos sistemas de recompensa e carreiras" (Chambers, 1994: 3-4). Esta conceção de profissionalismo colide com muitos dos valores proclamados pelas organizações do terceiro setor, que procuram "alternativas" ao "pensamento dominante" e demarcar-se das formas de gestão e administração dos setores lucrativo e público.

Defendendo uma nova ênfase em abordagens de gestão alternativas às de profissionalização de “cima para baixo”, como sejam as centradas no *empowerment*, participação e outras abordagens “de baixo para cima”, e partindo do trabalho de Chambers e outros autores, Korten identificou um conjunto de abordagens de gestão associadas ao “novo profissionalismo do desenvolvimento” (*new development professionalism*), no qual:

Ao invés de apoiarem o controlo central, [as ONG] apoiam a autoavaliação e autocorreção guiadas por uma forte orientação para o cliente do serviço e para um sentido de missão bem definido. Sistemas de gestão altamente desenvolvidos fornecem fluxos de informação ricos no sentido de facilitar estes processos de autogestão (Korten, 1987 *apud* Lewis, 2001: 22).

É no âmbito deste debate que se questiona até que ponto a gestão das OTS evoluirá no sentido de uma diversidade de estruturas e abordagens, ou se reverterá num processo de standardização tendente a um **isomorfismo** unidirecional destas organizações (DiMaggio e Powell, 1983, 1991). Os processos de profissionalização são frequentemente entendidos como burocráticos e criadores de barreiras entre as organizações e a comunidade, neutralizando formas alternativas de prestação de serviços (Carey, Annette e Barraket, 2009). No entender de Hwang e Powell (2009: 171), “o movimento generalizado de profissionalização do terceiro setor poderá ter o efeito, eventualmente inesperado, de transformar um grupo heterogéneo de organizações num setor distinto, coerente, com um conjunto de rotinas organizacionais comum”. Muitos argumentam que, com este processo, as OTS tornam-se menos distintas na sua natureza – os seus **valores são diluídos** e a base para a sua vantagem comparativa perde-se (Arvidson e Kara, 2013: 7).

Nesta linha, uma **potencial armadilha** para o sector que decorre da profissionalização das OTS pode ser a dominância de organizações institucionalmente fortes, com pessoal qualificado, que funcionam como organizações profissionais, em detrimento das organizações de caráter cívico e voluntário (Jailobaeva, 2014). David Lewis (2001) refere que o perigo atual para as OTS é que a profissionalização e as abordagens *mainstream* do desenvolvimento levem ao predomínio das orientações instrumentais em detrimento dos objetivos de expressão. É certo que a competição por recursos poderá não levar necessariamente ao isomorfismo, particularmente nas situações em que as entidades

financiadoras incentivam a inovação e a diversidade. No entender de alguns autores, no entanto, tem havido uma tendência geral por parte destas para anteporem a burocratização e eficiência à diversidade e inovação (Salamon, 1999; Edwards e Hulme 1995). Em todo o caso, a profissionalização do terceiro setor terá implicações significativas no seu núcleo identitário, enquanto arena para o voluntariado e participação, bem como para o seu desenvolvimento futuro e para o seu papel na sociedade em geral, especialmente considerando a presença do setor no cotidiano de cidadãos e cidadãs (Boris e Steuerle, 2006 *apud* Hwang e Powell, 2009: 270).

Apesar de associar a crescente profissionalização a uma série de riscos, a análise de Lester Salamon (2005), reportando-se ao contexto norte-americano, é otimista relativamente ao potencial da profissionalização das OTS. O autor faz um balanço positivo destas mudanças para o terceiro setor, associando-as ao fortalecimento da sua base financeira, à melhoria dos seus procedimentos, ao envolvimento de novos parceiros e recursos nas suas iniciativas e, em termos gerais, à melhoria da sua efetividade e da sua reputação (*ibidem*: 95-96). No mesmo sentido, David O. Renz (2016) conclui que não há razões para acreditar na ideia de que com a profissionalização da gestão do terceiro setor se perderá o seu carácter distinto na sociedade civil. No seu entender, se a educação e a profissionalização forem implementados de forma apropriada, o setor não perderá a sua orientação uma vez que o âmago de uma gestão efetiva deve estar ancorado no cumprimento da missão. Para ele o desafio é, em tempos difíceis, “não perder de vista a razão por que fazemos o que fazemos” (*ibidem*: 738).

De todo o modo, tem vindo a ser reconhecido que, nas últimas décadas, o terceiro setor se tem vindo a profissionalizar, transitando de uma “solidariedade informal” para um conjunto de atividades mais formalizadas, nomeadamente em resposta as pressões formais do Estado associadas à execução de programas públicos ou de interesse geral, às crescentes exigências ao nível da qualidade das respostas sociais, e à crescente concorrência de empresas, organizações públicas e organizações congéneres (Drori, Meyer e Hwang, 2006; Dobrai e Farkas, 2010). Configurando tendências de isomorfismo institucional com o mercado, ao longo da última década, em especial, tem-se assistido a movimentos vigorosos no sentido da importação de modelos e práticas das empresas, os

quais podem transformar grupos de solidariedade em **organizações mais instrumentais e racionais** (Hwang e Powell, 2009). Estes movimentos são fomentados e ampliados quer por gestores internos, que veem estas práticas como essenciais para o crescimento e sobrevivência das organizações, quer por “proselitistas externos” que incitam a difusão de ferramentas de gestão de negócios (*ibidem*: 271).

Numa visão otimista em relação a estes processos, Vernis *et al.* (2006: 148) defendem que se tem feito um progresso importante no sentido da melhoria da gestão das organizações do terceiro setor que deixaram de lado a cultura da “boa-vontade”, baseada no “fazer o bem”, para abraçar uma filosofia de trabalho que assente na noção de “fazer bem o bem”. No entender dos autores, esta mudança de perspetiva implica um salto qualitativo significativo para o desenvolvimento e fortalecimento do terceiro setor enquanto “terceiro pilar” da sociedade democrática.

Na sua obra, Lewis (2001: 198) conclui que há tantas armadilhas para as OTS que dependem dos elevados propósitos morais, boa vontade, trabalho árduo e senso comum, como para as que respondem acriticamente às pressões do ambiente para se profissionalizarem. Em última análise, trata-se de opções políticas (e não técnicas) que as OTS e os seus *stakeholders* têm que fazer. A forma como as OTS se relacionam com a questão da gestão depende do tipo de organização que querem ser, dos valores que pretendem expressar e da abordagem prosseguida para a prossecução da mudança social e económica (*ibidem*). A gestão nas OTS pode ser vista como tendo não só um **propósito instrumental** (i.e. alcançar os resultados desejados), mas também uma **qualidade expressiva**, isto é, um certo estilo ou cultura organizacional, apropriado quer para a tarefa em mãos quer para a preservação dos valores das pessoas envolvidas. O desafio é, portanto, **adotar modelos de gestão que melhorem o desempenho organizacional sem perder de vista os valores constitutivos das OTS** (Fowler, 1997; Lewis, 2001; Fowler e Malunga, 2010). Tal implica a adoção de uma **conceção holística da organização** (Anheier, 2005), que enfatiza a necessidade de se estabelecer uma relação mais rica com a envolvente externa e interna e implica a clarificação da visão e objetivos das OTS, harmonizando uma série de fatores internos e externos.

2.4. Abordagens à eficácia organizacional e desempenho das organizações do terceiro setor³⁷: contributos da teoria das organizações

O conceito³⁸ de eficácia organizacional é elusivo, contestado e particularmente difícil de apreender nas organizações do terceiro setor (Quinn e Rohrbaugh, 1983; Herman, 1990; Tasse, Murray e Cutt, 1998). A eficácia, enquanto **medida de “valor” ou “sucesso” organizacional**, vem atraindo há várias décadas a atenção de analistas no domínio das ciências sociais. Mais recentemente, a questão da eficácia passou a assumir relevância acrescida devido a exigências mais explícitas de prestação de contas, transparência e responsabilização financeira (Cutt e Murray, 2000; Murray, 2005; Ebrahim e Weisband 2007).

As revisões da literatura sobre eficácia organizacional tendem a enquadrá-la em campos distintos, mais ou menos comunicantes entre si. Uma das distinções mais comuns tem sido entre quem vê a eficácia como um processo pretensamente **racional**, e quem vê como um processo essencialmente **político** (Quinn e Rohrbaugh, 1983; Murray e Tassie, 1994; Herman e Renz, 2004). A primeira escola é representada sobretudo pelas **ciências empresariais positivistas** e por profissionais de avaliação. Nelas vigora o pressuposto de que os factos sobre o desempenho podem ser apurados, e, ainda que os julgamentos subsequentes possam ser falíveis, eles podem aproximar-se da “verdade”. Nas **escolas pós-**

³⁷ Ainda que alguns teóricos façam a distinção entre eficácia e desempenho (por exemplo, Liket, Garcia e Maas, 2014), neste trabalho os conceitos serão usados indistintamente, à semelhança do que acontece em muita da literatura sobre o tema (Kaplan, 2001; Henri, 2004; Baruch e Ramalho, 2006; Carton e Hofer, 2006). Reconhece-se, não obstante, que os dois conceitos radicam de tradições de investigação distintas: respetivamente, da teoria organizacional e da gestão estratégica. A eficácia organizacional surge na teoria organizacional sobretudo numa perspetiva de constructo (de tentativa de delimitação concetual), ao passo que o desempenho, em geral associado a sistemas de avaliação de desempenho, reflete mais uma perspetiva de processo (de quantificação da eficiência e eficácia através de um conjunto de métricas) (Henri, 2004). Na avaliação da eficácia são normalmente consideradas medidas de cariz mais preceptivo do que objetivo, enquanto na avaliação do desempenho, as medidas de natureza objetiva predominam (McCabe e Dutton, 1993, *apud* Ramalho e Baruch, 2006). O desempenho surge geralmente ligado a um conjunto vasto de dimensões, tais como crescimento organizacional e saúde financeira, enquanto a eficácia é particularmente associada ao alcance de objetivos organizacionais (Liket, Garcia e Maas, 2014).

³⁸ A condição concetual “caótica” (Cameron e Whetten, 1983) - falta de acordo entre os académicos relativamente a uma definição, quadro teórico, ou método de avaliação - da eficácia organizacional tem motivado diversos/as teóricos/as a designá-la de constructo (por exemplo, Henry, 2004; Cameron e Whetten, 1983; Campbell, 1977). Os constructos distinguem-se dos conceitos dado que os conceitos são abstrações de ocorrências que podem ser facilmente observadas ou medidas e os constructos inferências a um nível superior de abstração (Quinn e Rohrbaugh, 1983).

modernas da teoria organizacional domina uma percepção "política" da eficácia. No enquadramento desta perspectiva, mais em linha com o **paradigma construtivista**, acredita-se que quem avalia e quem é avaliado tem frequentemente objetivos e concepções de eficácia dissonantes, e que nem tudo pode ser medido objetivamente. Neste sentido, a avaliação é entendida como um processo em que as partes tentam negociar uma interpretação da eficácia e no qual as visões subjetivas do lado com maior poder tendem a dominar.

A literatura académica sobre a eficácia das OTS oferece uma grande variedade de modelos para explorar a eficácia organizacional, o que tem levado à afirmação de que existem tantos modelos como estudos de eficácia organizacional (Herman e Renz, 1999). Não obstante, nas suas revisões dessa literatura, Forbes (1998) e Lecy, Schmitz e Swedlung (2012) destacam três modelos principais que, isoladamente ou em combinação, sustentam uma série de outros modelos: o dos objetivos, o do controlo dos recursos (perspetiva dos recursos sistémicos), e o reputacional/socio-construtivista (de múltiplos constituintes). Martz (2013) acrescenta ainda um quarto modelo, que considero relevante integrar nesta revisão: o modelo dos processos.

Na **abordagem dos objetivos** à eficácia, os objetivos são geralmente aceites como parte da cultura, do desenho e da estrutura organizacional, sendo a eficácia associada ao grau em que a organização consegue atingi-los. A "gestão por objetivos" (Drucker, 1954), que considera que o critério essencial para avaliar o desempenho é determinar em que medida a organização executou as tarefas identificadas como necessárias, constitui o arquétipo do modelo orientado para os objetivos. Do mesmo modo, a pesquisa académica contemporânea em gestão do desempenho e planeamento estratégico tem reforçado a relação positiva entre definição e uso de metas e o desempenho organizacional (Martz, 2013).

A **abordagem dos recursos sistémicos** remete para a relação que se estabelece entre a organização e o seu ambiente, associando eficácia à capacidade da organização de assegurar recursos críticos da sua envolvente. Esta abordagem associa a realização de objetivos à sobrevivência organizacional, assumindo que as organizações que atingem os seus objetivos são também mais propensas a receber apoio financeiro continuamente

(Yuchtman and Seashore, 1967 *apud* Renz e Herman, 2016: 275). Na abordagem dos recursos sistémicos, contorna-se o desafio de identificar empiricamente o impacto enquanto medida do progresso face a objetivos específicos, focando a atenção na componente de sustentabilidade da eficácia organizacional (Lecy, Schmitz e Swedlung, 2012: 440).

A **abordagem reputacional/dos constituintes** baseia-se na agregação de medidas subjetivas de desempenho, conforme relatado por informantes-chave ou partes interessadas da organização. Esta escola enfatiza a *perceção* da organização como crucial para a compreensão da respetiva eficácia. A eficácia resulta, nesta abordagem, da capacidade que a organização tem para responder às diferentes aspirações e expetativas dos seus constituintes e ao que eles próprios assumem como critérios de eficiência e eficácia. Os critérios, muitas vezes conflituosos, são ponderados de acordo com o nível de controlo dos constituintes sobre os recursos da organização (Connolly, Conlon e Deutsh, 1980). Enquanto os/as construtivistas sociais visualizam a reputação como um facto social independente - e possivelmente mais importante - das medidas objetivas de eficácia, a perspetiva racionalista entende a reputação sobretudo como o registo do comportamento passado, o qual constitui uma das muitas componentes relevantes para a avaliação do desempenho de uma dada organização (Sharman, 2007: 35).

Por fim, a **abordagem de processos** baseia-se nos aspetos fundamentais dos modelos orientados por objetivos e sistemas, mas coloca uma forte ênfase nos processos internos da organização (Martz, 2013: 390). Nesta abordagem, tal como na dos objetivos, as metas são usadas para avaliar o desempenho organizacional, todavia, o foco não é o alcance dos objetivos mas sim na sua otimização (*ibidem*). Por outro lado, tal como na abordagem dos recursos sistémicos, também nesta os constrangimentos ambientais que limitam ou impedem que certas metas sejam plenamente alcançadas são reconhecidos e ponderados (*ibidem*). O modelo de processos, originalmente desenvolvido por Steers (1976 *apud* Martz, 2013: 389), fornece a quem avalia uma estrutura para analisar os principais processos que influenciam o desempenho. Embora não seja especificamente destinado a avaliar o desempenho organizacional, este modelo identifica práticas que são fundamentais para a implementação de processos efetivos e é frequentemente usado para

extrair conclusões avaliativas do desempenho organizacional do ponto de vista do processo. O *balanced scorecard*, introduzido na década de 1990 por Kaplan e Norton (1992), é ilustrativo desta abordagem³⁹. Trata-se de um sistema de medição de desempenho que foca os esforços de melhoria organizacional na melhoria do processo.⁴⁰ Outro exemplo da abordagem de processos é o modelo de “gestão baseada nos resultados”. Também introduzida na década de 1990, é uma estratégia de gestão ampla que se centra na forma como uma organização gere os seus processos e recursos para alcançar resultados específicos e mensuráveis⁴¹.

No quadro que se segue apresentam-se, de forma sintética, os principais pressupostos, asserções e críticas associados a cada uma das abordagens à (avaliação da) eficácia organizacional.

³⁹ Em algumas análises ferramenta é associada às abordagens híbridas/multidimensionais (por exemplo, Leczy, Schmitz e Swedlung, 2012).

⁴⁰ O *Balanced Scorecard* veio introduzir uma perspetiva mais ampla dos sistemas tradicionais de medição de desempenho utilizados nas organizações, englobando quatro dimensões operacionais que integram medidas financeiras, de relacionamento com clientes, de desempenho operacional, e de aprendizagem e crescimento. A inclusão de objetivos organizacionais, assim como de uma visão sistémica e da ênfase na aprendizagem e crescimento dos/as colaboradores/as, ressoam o modelo de processo da avaliação do desempenho organizacional avançado por Steers (Martz, 2013: 390). Apesar de as quatro perspetivas do *balanced scorecard* terem sido originalmente desenvolvidas para empresas, os seus criadores defendem que o modelo pode ser aplicado noutro tipo de organizações, com outra ordenação ou incluindo outras perspetivas. Para Kaplan e Norton (2004), há duas diferenças fundamentais entre o modelo para empresas e o aplicado às organizações do terceiro setor: a missão desloca-se para o topo do *balanced scorecard*, já que a aspiração última é a da concretização de uma missão criadora de valor social e não o desempenho, e a perspetiva do cliente passa a ser a mais importante. Paton (2003) foi mais longe e propôs uma abordagem *scorecard* especificamente para OTS, sugerindo que estas devem desenvolver e utilizar cinco categorias gerais de medidas: (a) medidas de desempenho atual de cada atividade ou programa na organização; (b) medidas de desempenho subjacente (de longo prazo) de cada atividade/programa, considerando formas alternativas de uso dos recursos; (c) medidas de solidez financeira e de gestão (abordando a questão de como está a ser conduzida a organização); (d) medidas de renovação ou de declínio (abordando a questão de saber se a organização está ou não a reforçar a sua capacitação); e (e) medidas de impacto social da intervenção da organização.

⁴¹ Embora os conceitos da “gestão baseada em resultados” e respetivas práticas para avaliar o desempenho organizacional possam remontar ao conceito de gestão por objetivos, o seu enfoque expande-se para os processos organizacionais (por exemplo, programação, mobilização de recursos e fornecimento de serviços) (Martz, 2013: 390).

Quadro 6. Pressupostos e críticas das abordagens à eficácia/desempenho organizacional

Modelos	Pressupostos e Atributos	Críticas
Objetivos	<p>As organizações são racionais, orientadas por objetivos;</p> <p>Os objetivos organizacionais são específicos, mensuráveis, realistas, operacionais, significativos e relevantes para o propósito da organização;</p> <p>Concentra-se nos efeitos resultantes da atividade da organização; Conduz a uma conclusão sumativa;</p> <p>O alcance das metas é o principal critério do desempenho organizacional</p>	<p>Existência de objetivos múltiplos, muitas vezes contraditórios/conflituantes;</p> <p>Dificuldades ao nível da medição de resultados;</p> <p>Preferência é dada a valores de determinadas pessoas e não de outras;</p> <p>Premissas de valor restrito;</p> <p>Os efeitos colaterais e os impactos não esperados não são identificados</p>
Recursos sistémicos	<p>As organizações são sistemas naturais ou abertos que interagem com o ambiente;</p> <p>Concentra-se no equilíbrio de recursos e sustentabilidade;</p> <p>Focada em medidas de sobrevivência e crescimento organizacional;</p> <p>Considera o desempenho dos subsistemas que contribuem para o desempenho organizacional;</p> <p>A aquisição de recursos escassos e valorizados sugere uma organização eficaz</p>	<p>Definição restrita de eficácia, que não chega endereçar ao desempenho;</p> <p>Não aborda forças externas que escapam ao controlo da gestão e que podem afetar o desempenho da organização (por exemplo, mudanças socioculturais, legislativas ou económicas);</p> <p>Potencial negligência de públicos-alvo ou <i>stakeholders</i> chave quando estes não são explicitamente parte de um subsistema</p>
Reputacional	<p>As organizações são sistemas naturais ou abertos onde os participantes com diferentes graus de poder disputam o controlo sobre os recursos;</p> <p>Uma organização eficaz é aquela que satisfaz os requisitos dos seus constituintes;</p> <p>Concentra-se nas preocupações daqueles que podem mais impactar ou garantir a sobrevivência da organização;</p> <p>Promove a participação e a legitimidade organizacional</p>	<p>Desafio de pesar pontos de vista divergentes (de diversos <i>stakeholders</i>) sobre o que faz uma organização eficaz;</p> <p>Tende a favorecer as ligações mais poderosas dentro do ambiente organizacional;</p> <p>Uma organização pode ser considerada de eficaz, mesmo quando não possui uma vantagem competitiva</p>
Processos	<p>Podem ser estabelecidas ligações causais entre os processos internos e os resultados organizacionais;</p> <p>Reconhece os constrangimentos ambientais e o seu impacto na consecução dos objetivos;</p> <p>A otimização de objetivos é o principal critério do desempenho organizacional</p>	<p>Não aborda especificamente as forças externas além do controle de gestão;</p> <p>Risco de a ênfase na eficiência interna dos processos pode distrair as organizações do reconhecimento e da ação sobre as mudanças macroambientais que tornam a organização pouco competitiva ou sua finalidade irrelevante</p>

Fonte: Elaboração da autora, com base, em especial, em: Forbes, 1998; Herman e Renz, 1999; Lecy, Schmitz e Swedlung, 2012 e Martz, 2013.

Em resposta aos desafios colocados por cada uma destas escolas, a literatura direcionou-se para a adoção de **modelos de eficácia mais complexos**, que normalmente incorporam aspetos das abordagens por objetivos, de recursos sistémicos, de múltiplos constituintes e de processos. **A ideia de que as medidas unidimensionais de eficácia não são úteis reúne amplo consenso na literatura, ainda que as medidas únicas ainda sejam prevaletentes na prática de muitas organizações do terceiro setor** (Lecy, Schmitz e Swedlung, 2012; Raposo, 2011). O consenso relativamente ao caráter multidimensional da eficácia, embora não conduzindo à emergência de um sistema de avaliação/medição da eficácia amplamente partilhado (Bradshaw, Murray e Wolpin, 1992), faz sobressair dois tipos de abordagens (Lecy, Schmitz e Swedlung, 2012; Herman e Renz, 1999; Forbes, 1998): as **abordagens reputacionais** (ou socio-construtivistas/"emergentes"/de múltiplos constituintes) e as **abordagens híbridas multidimensionais**.

As **abordagens emergentes** enfatizam a relação dialética entre os *stakeholders* e as estruturas, colocando a ênfase na compreensão das interações que têm lugar dentro e entre organizações, bem como nos papéis que desempenham a informação e a comunicação na formação de juízos sobre eficácia (Forbes, 1998: 195). Segundo esta perspetiva, ainda que a eficácia seja socialmente construída pelas ações e interações dos *stakeholders*, as suas implicações e consequências no financiamento e oportunidades das OTS são reais. As medidas de reputação afirmam o caráter flexível da eficácia e a importância de ter em consideração as perceções sobre eficácia por parte de múltiplos *stakeholders*. Forbes observa que a importância teórica dos estudos construtivistas "não deriva da sua capacidade de representar efetivamente a eficácia organizacional, mas sim do seu potencial para iluminar a forma como a eficácia é concebida, negociada e medida" (1998: 196). O desafio que ela coloca, porém, e que tem merecido crescente apoio, é desenvolver um modelo de eficácia adequado que capte a diversidade de interesses e pontos de vista presentes nas OTS.

A **abordagem híbrida** tenta integrar múltiplas medidas de eficácia num único modelo. Em geral, esta perspetiva integra elementos do modelo dos objetivos, o modelo dos recursos sistémicos e o modelo dos múltiplos constituintes. O "modelo multidimensional e integrado de eficácia das organizações não lucrativas" (MIMNOE) de

Sowa, Seldon e Sandfort (2004)⁴²; e o modelo de contingência para a medição dos resultados organizacionais, desenhado por Ebrahim e Rangan (2010)⁴³, são exemplos relevantes do tipo de abordagem multidimensional.

As abordagens de avaliação da eficácia tendem a focar-se no estabelecimento de padrões e **linhas de orientação para as organizações, individualmente**, e menos no setor como um todo. A diversidade de papéis e das funções do setor e com as discussões conceituais e empíricas acerca das suas fronteiras, aduzem dificuldades nas análises da determinação do contributo ou do **impacto global do setor** em contextos particulares (Harlock, 2013: 6), ainda que já se tenham feito algumas tentativas relevantes. Uma das iniciativas mais sonantes no sentido de desenvolver critérios genéricos para medir a eficácia foi apresentada por Kendall e Knapp's (2000). Os autores propuseram um modelo com oito domínios de desempenho (economia, eficácia, escolha/pluralismo, eficiência, equidade, participação, advocacia e inovação), sugerindo um conjunto de indicadores no sentido de abordar e apreender os vários papéis e contribuições políticas, económicas e sociais do setor numa variedade de contextos. Um outro exercício relevante encontra-se nos resultados do projeto de investigação "Third Setor Impact", coordenado por Bernard Enjolras.⁴⁴ É proposto um modelo de impacto assente numa série de domínios de impacto

⁴² Este modelo destaca dois níveis ou dimensões distintas de eficácia: eficácia da gestão e eficácia de programas. Ambas as dimensões são decompostas em duas subcomponentes: capacidade e resultados. Uma das singularidades deste modelo deriva dos tipos de indicadores associados a cada uma das subdimensões da eficácia da gestão e dos programas: indicadores objetivos e indicadores percetivos. Para ilustrar como o modelo pode ser usado empiricamente, as autoras propõem um método de análise que explora as inter-relações entre as múltiplas dimensões que o compõem.

⁴³ O modelo da contingência é arrojado nos seus pressupostos e asserções. A sua essência – a contingência – sugere, desde logo, que não existem panaceias para medição de resultados em contextos sociais complexos. O enquadramento proposto baseia-se na noção de que "o que se deve medir depende do que se procura alcançar", e é construído em torno de duas dimensões: a teoria da mudança da intervenção (a lógica causal subjacente à intervenção) e a sua estratégia de operacionalização (como é que a organização implementa intervenções que promovem a sua missão). Os autores usam este modelo para definir quatro tipos de resultados (resultados "compartimentados" - *niche results* -, resultados integrados, resultados institucionais e resultados de ecossistema), e o tipo de medidas adequadas a cada um deles. Para os autores, medir o impacto apenas faz sentido num conjunto limitado de circunstâncias: quando uma organização opera ao nível do "ecossistema", e ainda assim pode exercer um controlo suficiente sobre os resultados para atribuir os impactos ao seu trabalho. Esta abordagem propõe que, dada a diversidade do trabalho, objetivos e capacidades das organizações do setor social, algumas organizações devem medir impactos de longo prazo, enquanto outras devem ficar-se pela mensuração de resultados de curto prazo.

⁴⁴ O projeto visou compreender o alcance e a escala do terceiro setor na Europa, o seu impacto atual e potencial, e as barreiras que impedem o terceiro setor de contribuir plenamente para o bem-estar do continente europeu (<http://thirdsectorimpact.eu/>).

(bem-estar e qualidade de vida, inovação, envolvimento cívico, advocacia, desenvolvimento das comunidades locais, impactos económicos e recursos humanos) que podem ocorrer numa perspetiva de curto, médio e longo prazo, e a nível micro, meso ou macro e que deverão ser medidos a partir de um “conjunto de indicadores consensualizados” (Enjolras, 2015). Argumentando a necessidade de identificar critérios multidimensionais para captar o impacto do setor, o autor chama a atenção para os desafios que se colocam à medição, alertando para o facto de as avaliações quantitativas poderem ser “incompletas, tendenciosas e, portanto, falaciosas”, ao concentrarem-se nos resultados facilmente mensuráveis, ignorando os de difícil medição (Enjolras, 2016: 2).

Há que reconhecer a relevância dos exercícios teóricos que procuram estabelecer enquadramentos ou modelos para a avaliação, já que poderão constituir recursos importantes para apoiar a capacitação das organizações neste domínio (Hall *et al.*, 2003; Ellis e Gregory, 2008; Ógáin, Lumley e Pritchard, 2012). **O estabelecimento de um modelo de avaliação comum para o setor é, no entanto, desafiado por uma série de fatores.** Como notam Cameron e Whetton (1983), as decisões sobre o domínio da atividade a avaliar, sobre os *stakeholders* a incluir na avaliação, sobre o nível de análise, e muitos outros fatores dificultam o estabelecimento de um padrão comum. Por outro lado, a afirmação de uma abordagem “one-size-fits-all” à avaliação do desempenho das OTS é desafiada pela diversidade do setor – dos seus papéis, interesses, atividades e objetivos – e pela complexidade dos ambientes nos quais operam. Além disso, as OTS têm múltiplos *stakeholders* – entidades financiadoras, filantropos/as, utilizadores/as e beneficiários/as dos serviços, trabalhadores/as, voluntários/as e a comunidade local ou população mais alargada –, que poderão ter diferentes expectativas e perceções sobre o papel e sobre os resultados das OTS, assim como capacidades e oportunidades diversas para fazer exigências explícitas sobre o tipo de informação de avaliação a valorizar (Kendall e Knapp, 2000: 107; Harlock, 2013: 6). Daqui decorrerão múltiplos e variados imperativos de recolha de informação avaliativa por parte de *stakeholders* externos e por parte das próprias organizações, assim como diferentes expectativas sobre os tipos de dados que são mais úteis. Esta pluralidade faz com que seja difícil encontrar um quadro de avaliação comum que vá ao encontro de todos os interesses e necessidades de todas as partes interessadas.

As hesitações a respeito da pertinência de um modelo de avaliação que constitua uma solução "única" adaptada a todas as situações, e as discordâncias sobre qual a linguagem adequada para apreender os resultados dos projetos e os impactos mais amplos na comunidade do terceiro setor, suscitam divergências (Arvidson, 2009: 9).

Estes fatores têm sustentado a ideia, defendida em diversas análises (Forbes, 1998; Herman e Renz, 1999; Sowa *et al.*, 2004; Carman, 2007; Eckerd e Moulton, 2010; Ógáin, Lumley e Pritchard, 2012; Millar e Hall, 2013), e mesmo por organismos internacionais⁴⁵, de que as **abordagens diferenciadas e multidimensionais são mais adequadas do que as uniformizadas**, e de que diferentes organizações poderão beneficiar de diferentes práticas. Todavia, nas OTS existe o receio de que a preocupação com a profissionalização do setor e o impulso para a avaliação/medição do desempenho e prestação de contas vá comprometer essa heterogeneidade, resultando na promoção e adoção de práticas de avaliação que servem pouca utilidade para muitas organizações sem fins lucrativos, por vezes à custa de avultados gastos (Eckerd e Moulton, 2010; Simister e Smith, 2010).

2.4.1. A avaliação do desempenho e do impacto do terceiro setor enquanto processo socialmente construído

Alguns estudos produzidos nos últimos anos (Arvidson, 2009; Lyon e Arvidson, 2011; Hall, 2012; Arvidson e Lyon, 2013; Harlock, 2013), têm chamado a atenção para a necessidade de ir além das **abordagens normativas** de muita da pesquisa existente sobre a medição e avaliação do desempenho/impacto no TS, e de conduzir uma análise mais crítica, tanto das implicações como dos desafios que se colocam à avaliação nas OTS. Ao analisar criticamente as lógicas subjacentes aos processos de avaliação e medição nas OTS, estas análises **têm contestado os pressupostos acerca da neutralidade das avaliações** (Hall, 2012; Barman, 2007; Arvidson, 2009) e têm destacado alguma contestação sobre a

⁴⁵ O sub-grupo para a medição de impacto do Expert Group on Social Business (GECES) da Comissão Europeia, na sua publicação "Proposed Approaches to Social Impact Measurement", reconhece a inadequação da criação de um "conjunto rígido de indicadores" numa abordagem "top-down e "one-size-fits-all" de medição do impacto social, devido, entre outros fatores, à "variedade do impacto social perseguido pelas empresas sociais" e à "grande variedade na natureza e objetivos das atividades" destas organizações (CE, 2012: 6-7).

forma como a eficácia (o impacto social, em particular) das OTS pode ser objeto de mensuração (Westall, 2009; Lyon e Arvidson, 2011; Arvidson e Kara, 2013; Harlock, 2013).

Leituras mais críticas dos conceitos e práticas de medição e avaliação do desempenho/impacto (Barman, 2007; Arvidson, 2009; Nicholls, 2009; Hall, 2012; Harlock 2013; Manetti, 2014) argumentam que as avaliações não devem ser entendidas apenas como exercícios técnicos e científicos, antes, revelam e **refletem um conjunto de valores, premissas e ideais sobre o que deve ser medido e como**. Tais valores, premissas e ideais variam ao longo do tempo e em diferentes contextos sociais e, portanto, as avaliações e os seus resultados devem ser tratados como atividades não inteiramente objetivas, mas sim entendidas e consideradas no contexto particular em que são realizadas (Harlock, 2013: 9). Carmel e Harlock (2008), por exemplo, defendem que os objetivos e metas de desempenho estabelecidas para as OTS que prestam serviços públicos muitas vezes refletem as agendas de seus financiadores e do governo, priorizando atividades e formas de trabalho particulares com base na prática do setor privado ou empresarial, a fim de promover a eficiência e custo-eficácia, em prejuízo de outros objetivos.

Estas perspectivas promovem o questionamento de como diferentes ferramentas e práticas de medição e avaliação do desempenho utilizam e priorizam diferentes tipos de dados, informações e ideais sobre a criação de conhecimento. Um desacordo comum nos debates sobre as abordagens/metodologias mais adequadas para avaliar o desempenho e diz respeito ao argumento de que os estudos de caso e as histórias podem conter subjetividade, enquanto os indicadores de desempenho e as estatísticas são mais objetivos e fiáveis (Jacobs, Barnett e Ponsford, 2010; Hall, 2012). Tais desacordos refletirão uma crença normativa na superioridade de determinadas abordagens de avaliação relativamente a outras, e não as forças e fraquezas inerentes a qualquer metodologia/técnica particular.

Os debates políticos e académicos mais recentes sobre o desempenho do terceiro setor têm girado em torno da apreensão e demonstração do **impacto social**⁴⁶ e/ou valor

⁴⁶ Neste estudo “impacto social” é entendido por mudanças externas mais amplas para a sociedade, a economia e/ou o ambiente que os OTS podem gerar através das suas atividades (não se centrando apenas nos resultados diretos para beneficiários/as ou *stakeholders* individuais ou privados) (Arvidson *et al.* 2013).

social⁴⁷ do terceiro setor e das suas atividades (e.g., Bassi, 2015). Embora o impacto social tenha vindo a ser aludido como correspondendo aos benefícios mais amplos da atividade das OTS, em contraste com o impacto do sector mais estritamente na economia e/ou sistemas económicos particulares (Westall, 2009; Harlock, 2013), os debates sobre o valor social criado pelas atividades do TS têm sido fortemente influenciados por conceções económicas de valor. Westall (2009) defende que o **valor monetário e os resultados monetizáveis tenderam a dominar conceitos de valor, deixando pouco espaço para entendimentos alternativos.**

Esta abordagem reflete-se particularmente no desenvolvimento de métricas como o Retorno Social do Investimento (SROI), ou a necessidade de evidenciar o "valor acrescentado" (frequentemente em comparação com o sector privado). Tal viés é compreensível à luz das necessidades e agendas de determinados atores, particularmente dos governos, instituições financeiras e financiadores (Westall, 2009: 6). Enquanto para o Estado a preocupação essencial é assegurar o melhor uso do dinheiro público para atingir os objetivos de interesse público, para instituições financeiras e entidades financiadoras o valor criado em relação ao montante investido é uma preocupação central (*ibidem*). Como resultado dessa pressão e da natureza das métricas dominantes, quem financia o setor torna-se mais prescritivo e diretivo, induzindo ou forçando as OTS a aderir a práticas e ideias oriundas do mundo económico e empresarial, as quais, ironicamente, estão a ser contestadas e alteradas (*ibidem*).

Por outro lado, continuam a ser levantadas questões sobre se as medidas financeiras por si só são adequadas para captar e medir o desempenho e o impacto social do terceiro setor. Alguns/mas analistas (por exemplo, Westall, 2009; Arvidson *et al.*, 2010, 2013; Lyon e Arvidson, 2011) observam que as ferramentas que incidem sobre os impactos externos e *proxies* financeiras, tais como SROI: representam desafios para captar e avaliar a experiência qualitativa e o conhecimento tácito, que é mais difícil de expressar em

⁴⁷ O conceito de valor social, central na literatura e prática do empreendedorismo social, refere-se ao "resultado de um conjunto de respostas e efeitos de atividades, produtos ou serviços que beneficiam e são valorizados por pessoas cujas necessidades não estão a ser supridas por outros atores, nomeadamente pelo Estado e pelo mercado. O resultado tem valor social na medida em que 'é benéfico porque é valorizado'" (Young, 2006 *apud* Parente, 2014: 243).

medida; negligenciam outras dimensões da criação de valor diferente do valor monetário; negligenciam a importância dos valores organizacionais e culturais implicados no processo de medição do impacto; e oferecem oportunidades de discricionariedade no processo de avaliação; na escolha de quem realiza avaliações, na seleção e identificação de indicadores, na recolha, análise de dados e na apresentação dos resultados.

Estas perspetivas, enquadráveis no **paradigma transformativo**, vêm sublinhar o **caráter político da avaliação**, no sentido em que ela é sempre contextualizada na promoção de diferentes valores e posições políticas (Greene, 1994; Chouinard, 2013) e em que a escolha de uma abordagem de avaliação tem implicações políticas e normativas (Rossi *et al.*, 2004). A avaliação pode, nesta linha, acrescentar valor a quem é marginalizado, mas também incorporar e reproduzir as desigualdades e as relações de poder (Arvidson e Kara, 2013: 15). A avaliação não aprecia apenas as intervenções mas é também promotora de valores políticos e metodológicos (*ibidem*: 15). Por exemplo, uma avaliação que vise promover a equidade no acesso à educação terá um foco diferente daquela que prioriza a eficiência na prestação de serviços.

2.5. Relações de accountability e avaliação nas organizações do terceiro setor

Reenviando às suas bases epistemológicas⁴⁸, a avaliação tem sido tradicionalmente retratada como um instrumento racional que quem gere as OTS pode usar para recolher informações necessárias para tomar decisões sobre como melhorar as intervenções e atingir as metas organizacionais.

O objetivo afirmado e ideia por detrás da avaliação estão frequentemente enraizados nos pressupostos de racionalidade e de causalidade. Muitos dos modelos de avaliação e pressupostos dos usos da avaliação derivam desta lógica (Højlund, 2014), que está ligada a entendimentos realistas e positivistas comumente encontrados no pensamento económico clássico e na escolha racional. A teoria da escolha racional baseia-se e na ideia de que os indivíduos e as organizações vão "fazer escolhas fundamentadas sobre a desejabilidade de adotar diferentes cursos de ação" (Dunn, 1981: 226, *apud*

⁴⁸ As origens da avaliação, dominadas por premissas positivistas, situam-na como inerentemente racionalista, causal e evolucionária (Sanderson, 2010).

Carman, 2011: 351). O enfoque no valor racional, objetivo e nos aspetos técnicos da avaliação (com acento no seu uso instrumental), de muita da literatura sobre avaliação fundamenta-se implicitamente nesta teoria (*ibidem*: 351). A organização ideal seria auto avaliativa, monitorizaria continuamente as suas atividades para determinar até que ponto estaria a cumprir os seus objetivos, ou mesmo se esses objetivos seriam adequados (*ibidem*: 351). O pressuposto fundamental das relações lineares de causa e efeito está intrinsecamente ligado à categoria de uso instrumental: a avaliação melhorará o desempenho da organização através da aplicação instrumental dos seus resultados (conclusões e recomendações).

Alicerçando-se nestes pressupostos, quem escreve sobre avaliação e advoga o seu uso emprega frequentemente a perspetiva da escolha racional, assumindo que quem gere as organizações vai optar por fazer a avaliação, uma vez que se entende que ela poderá ajudar a tomar decisões mais informadas. O resultado tem sido a proliferação – sobretudo nos países anglo-saxónicos – de instrumentos normativos e descritivos concebidos para tornar a avaliação e medição de desempenho mais acessível à gestão dos setores público e não lucrativo (por exemplo, Moseley e Dessinger, 2010; W.K. Kellogg Foundation, 2010; Wholey, Hatry e Newcomer, 2010; Poister, Aristigueta e Hall, 2014).

No entanto, diversas análises, particularmente no campo da avaliação de programas e políticas, notam que a avaliação não tem cumprido a sua promessa de informar quem toma decisões (Patton, 1997; Alkin e Taut, 2003; Shulha e Cousins, 1997). Por outro lado, ao contrário do principal pressuposto racional da finalidade da avaliação, as organizações ou atores utilizam a avaliação com fins muito diversos, designadamente para satisfazerem necessidades/interesses de quem as financia e outros *stakeholders*, demonstrarem competência, alcançarem legitimidade e obterem financiamento (por exemplo, Lyon e Arvidson, 2011; Arvidson e Lyon, 2014; Ógáin *et al.*, 2012; Mitchell e Berlan, 2016). A teoria da escolha racional é, pois, insuficiente para compreender o comportamento das organizações no que toca à adoção de práticas de avaliação e aos usos ou não-usos dos seus resultados. Com efeito, existem outras perspetivas teóricas, que levam em linha de conta as pressões ambientais e institucionais, e que podem fornecer

contributos adicionais para entender porque é que as OTS fazem avaliação e como é que elas usam os resultados.

Nestas abordagens, a avaliação deixa de ser concebida como um instrumento eminentemente técnico e racional, e passa a ser vista numa perspetiva relacional, passando também a ser concebida como um **mecanismo de prestação de contas/responsabilização**.

Embora o conceito de *accountability* (ou prestação de contas) continue a ser um conceito "abstrato e composto", "oferecendo uma gama de significados possíveis" (Ebrahim, 2003b: 193), em essência, prestação de contas significa ser responsável perante determinada parte, por um desempenho esperado (Bies, 2010: 1061). Partindo desta perspetiva genérica, é possível classificar em **dois grandes grupos as definições de accountability**: as definições estritas⁴⁹ e as definições amplas (Kearns, 1996: 41). A **visão estrita** centra-se na capacidade de dar resposta a uma autoridade superior, em termos do cumprimento dos mandatos explícitos e remete para uma responsabilização "legal". A **visão ampla**, por seu turno, remete para o escrutínio por parte do público em geral e reporta-se a uma responsabilização "moral", sendo que os padrões podem ser implícitos, subjetivos e estarem em constante mudança. A **responsabilização legal** ocorre quando as partes aceitam formal e oficialmente a responsabilidade e se comprometem com alguma forma explícita de reporte. A **responsabilização moral**, que Leat (1988 *apud* Lewis, 2001) também designa de **responsiva**, ocorre quando o reporte não é exigido legalmente, mas as partes acreditam que existe uma obrigação de responsabilização perante outros/as.

Procurando expandir o conceito de prestação de contas, Santos (1999: 10) e Edwards e Hulme (1996: 967) distinguem **responsabilização ascendente** de **responsabilização descendente**. Enquanto a responsabilização ascendente é feita a entidades financiadoras e agências reguladoras, por exemplo, organismos de tutela, e tem sobretudo um enfoque no "emprego de verbas designadas para propósitos designados (Najam, 1996: 342), a responsabilização descendente é feita a membros, utilizadores e comunidade em geral, referindo-se primordialmente a "grupos aos quais as ONG prestam serviços", podendo incluir também "comunidades e regiões indiretamente afetadas pelos

⁴⁹ Exemplo de uma definição ancorada numa perspetiva estrita é a avançada por Edwards e Hulme (1996: 967, 1995: 9) que definem *accountability* como "o meio através do qual indivíduos e organizações reportam a autoridade(s) reconhecida(s) e são responsabilizados pelas suas ações".

programas das ONG” (Najam, 1996: 345). Najam (1996) destaca ainda uma terceira categoria, a de **responsabilização perante si próprias**, que respeita à responsabilidade das OTS perante a sua missão e pessoal, incluindo decisores/as e pessoal do terreno. Nesta mesma linha, Cornwall, Lucas e Pasteur (2000: 3) fazem a distinção entre “ser responsabilizado por” e “assumir responsabilidade”. Enquanto esta última dimensão, que Fry (1995) designa de “responsabilidade sentida”, reflete um estado intrínseco e subjetivo que espelha uma aceitação profunda da responsabilidade, constituindo a vertente interna do comportamento responsável, a primeira constitui a parte externa do comportamento responsável, e resulta da dinâmica social pela qual se é impelido a prestar contas a outrem, por desvios entre a ação pela qual se é responsável e as normas e expectativas relevantes. Estas perspetivas acrescentam à **dimensão externa** da responsabilização (a obrigação de cumprir as normas prescritas de comportamento) uma **dimensão interna**, expressa por meio de ações individuais e da missão organizacional (Ebrahim, 2010: 3).

Outra distinção útil e feita por Edwards e Hulme (1995: 9-11) e Najam (1996: 351), entre **prestação de contas funcional** e a **prestação de contas estratégica**. A primeira é centrada no uso dos recursos, na visibilização de resultados a curto prazo (*outputs*) e na relação com os principais *stakeholders*; a segunda é centrada nos impactos mais amplos (na qualidade de vida das pessoas) e na relação com todos os *stakeholders* (incluindo parceiros e cidadãos/ãs).

No quadro que a seguir se apresentam-se, esquematizam-se as principais características associadas aos diversos tipos de prestação de contas no que toca à sua indução (interna ou externa), aos principais atores perante quem as OTS se “responsabilizam”, e ao tipo de expectativas a que respondem, em termos de conteúdos.

Quadro 7. Características dos tipos de prestação de contas

	Legal Explanatória ou c/ sanções Funcional	Moral Responsiva Estratégica	
	Ascendente	Interna	Descendente
Impulso	Externo (requisitos legais, ou de financiamento)	Interno (melhorar o desempenho na prossecução da missão)	Interno (valores organizacionais) / Externo (requisitos de financiamento)
Stakeholders	Entidades financiadoras/tutelares	Membros (pessoal decisor, trabalhador e voluntário)	Beneficiários/as, parceiros, comunidade, cidadãos/ãs
Conteúdo	Relação entre <i>inputs</i> e <i>outputs</i>	Missão	Missão/Impacto

Fonte: Elaboração da autora.

Como salienta Ebrahim (2005: 58-59), as organizações estão envolvidas num processo complexo e contínuo de equilíbrio entre prestações de contas que são conduzidas externamente (isto é, de cima para baixo, ou punitivas) e aquelas que são geradas internamente. A prestação de contas pode então ser definida como “os meios através dos quais indivíduos e organizações são compelidos externamente a justificar as suas ações e como os meios através dos quais eles assumem internamente a responsabilidade de dar forma e escrutinar continuamente a missão, metas e desempenho organizacionais” (Ebrahim, 2003b: 194).

2.5.1. Perspetivas teóricas sobre a adoção de sistemas de *accountability*

O exame da literatura permite-nos perscrutar diversas perspetivas e teorias, em grande medida influenciadas pelas abordagens económicas e organizacionais, que enquadram e explicam a adoção de sistemas de *accountability* – e, mais concretamente, de avaliação e medição do desempenho nas OTS. Estas teorias incluem a teoria da agência, e a teoria dos *stakeholders*, a teoria da *stewardship*/administração, a teoria da dependência de recursos e a teoria institucional. Enquanto a teoria da escolha racional pressupõe que as organizações, tal como os indivíduos, se orientam para o alcance de objetivos, usando a informação para fazer escolhas deliberadas e racionais, estas teorias partem de pressupostos diferentes sobre o modo como as organizações operam, levando em consideração os efeitos do ambiente e a natureza das relações interorganizacionais.

Talvez o contributo mais significativo para o enquadramento da avaliação seja dado pela **teoria da agência ou teoria do *principal-agente*** (*agency theory/principal-agent theory*). Esta teoria baseia-se na observação de que alguns indivíduos (*principais*) procuram que as suas agendas sejam conduzidas por *agentes*. No âmbito desta perspetiva, a responsabilização é definida como o direito do *principal* pedir contas ao *agente* e de impor sanções se as contas prestadas forem inadequadas.

Esta abordagem postula que entidades financiadoras e OTS têm objetivos, interesses e motivações inerentemente diferentes. Como tal, o requerimento às organizações no sentido de reportarem elementos sobre resultados e outras informações sobre o seu desempenho constitui simplesmente parte do processo de monitorização inerente à relação entre o *principal* (financiador) e o *agente* (OTS). O ator atuando como *principal* visa moderar a gestão e o funcionamento organizacional do *agente*. A preocupação central da teoria da agência incide, pois, sobre a forma de garantir que o *agente* opera efetivamente em linha com os interesses do *principal*. De acordo com esta teoria, é provável que as organizações mascarem as suas realizações perante as entidades financiadoras, a fim de acederem a um financiamento ou contrato (seleção adversa) e que tendam a esquivar-se e não empenhar o nível de esforço acordado (risco moral) (Jensen e Meckling, 1976 *apud* Carman, 2010). A consequência é que tanto entidades financiadoras como OTS gastarão tempo e recursos para mitigar estes problemas através da implementação de sistemas de monitorização, medição de desempenho e supervisão.

Apesar do potencial importante para a explicação dos processos de *accountability*, a perspetiva do *principal-agente* tem recebido algumas críticas (por exemplo, McDonald, 1997; Ebrahim, 2005, 2009; Bies, 2010; Carman, 2010) que põem a claro algumas das suas limitações na explicação da introdução de práticas de avaliação nas organizações. Uma das críticas mais consistentes é apresentada por Ebrahim (2005: 196-199, 2009: 890-891). Segundo o autor, a perspetiva *agente-principal* contribui para a compreensão da prestação de contas, focando nas relações entre atores (*principais* e *agentes*) e nas estratégias utilizadas pelos *principais* para verem cumpridas as suas agendas. Fica, no entanto, aquém na abordagem dos problemas de *accountability* decorrentes da existência de múltiplos *principais*, da incongruência de interesses entre *principais* e *agentes*, dos mecanismos

internos da responsabilização, dos papéis dos *principais* na formatação do desempenho do *agente*.

Em primeiro lugar, as OTS são, frequentemente, responsáveis perante múltiplos *principais*, esperando-se que respondam aos interesses de membros, clientes e de entidades financiadoras e tutelares (Lindenberg e Bryant, 2001; Najam, 1996). O segundo problema está relacionado com este, surgindo quando a distinção entre *agente* e *principal* é ambígua, especialmente quando os interesses dos atores não são congruentes ou quando as entidades do terceiro setor trabalham com múltiplos *principais*. Por exemplo, numa relação entre um organismo do Estado e uma OTS, a instituição do Estado geralmente atua como *principal*, enquanto reguladora e financiadora das OTS. Mas a OTS pode ver-se como *agente* das pessoas e comunidades que serve, ao invés de *agente* do seu doador. Sem dúvida, a incongruência de interesses nem sempre é reconhecida pelos *principais* (como sejam doadores), que por vezes supõem que estão concordantes apenas porque o *agente* não tem no lucro uma motivação (Bogart, 1995: 164-165).

Um terceiro problema apontado por Ebrahim (2005) remete para a distinção já mencionada entre a dimensão externa da responsabilização e a “responsabilidade sentida”, de que decorre a importância de compreender “se a responsabilização é experienciada como um processo de controlo ou de capacitação” (Fry, 1995: 186). No entender do autor, *policy makers* públicos e financiadores/as das OTS tendem a concentrar-se em meios externos de assegurar a prestação de contas, como sejam medidas de monitorização, controlo e regulamentação. Ao fazê-lo, negligenciam as questões de responsabilização que são internas às organizações e aos indivíduos, tais como a integridade e a missão, e através dos quais *principais* e *agentes* atuam para o desenvolvimento de interesses e compromissos congruentes (*ibidem*: 182).

Uma quarta limitação da teoria da agência prende-se com o seu enfoque nos requisitos comportamentais dos *agentes*, enquanto se retira a ênfase dos *principais* que são, na realidade, responsáveis por influenciar o desempenho dos seus *agentes* através de uma série de meios. Como ilustra Catherine McDonald (1997) no seu estudo, os *principais* desempenham frequentemente múltiplos papéis em relação aos seus *agentes*. Quem gere um programa de financiamento pode ser responsável simultaneamente pela monitorização

do desempenho de uma entidade beneficiária e por lhe dar apoio na construção de capacidade de implementação e gestão financeira do projeto. Ademais, como também ressalta McDonald (1997), as relações entre *principais* e *agentes* são afetadas por processos políticos que têm impacto sobre o principal e nas relações entre *principal* e *agente*.

As **perspetivas dos stakeholders/partes interessadas** presentes na literatura sobre comportamento organizacional vêm preencher a lacuna desta e de outras abordagens sobre responsabilização, da falta de um olhar integrado sobre a forma como as organizações lidam com exigências múltiplas e por vezes concorrentes de prestação de contas. Grande parte do trabalho inicial neste domínio é creditado à obra de Edward Freeman (1984) sobre a “abordagem dos *stakeholders*” à gestão estratégica, na qual os *stakeholders* são definidos no sentido de incluir não só acionistas mas também outros indivíduos e grupos que podem afetar ou são afetados por determinado negócio. De forma similar, alguns/mas autores/as têm feito notar que também as OTS são responsáveis perante múltiplos atores. Embora os interesses de cada um destes principais possa coincidir até determinado ponto, potenciais conflitos ou ambiguidades nos seus interesses podem levar as OTS a alterar constantemente a sua retórica para agradar a múltiplos *principais* ou dogmaticamente afirmar as suas próprias missões. Além disso, é frequente as OTS resistirem a ver-se a si próprias como *agentes* de uma organização externa, vendo-se antes como sendo conduzidas pela sua missão. As OTS têm, pois, múltiplos *stakeholders* e são “responsáveis” de formas complexas e diversas perante uma variedade de grupos e interesses.

A **teoria da *stewardship***, contrastando com a da agência, sugere que financiadores e organizações do terceiro setor têm, na realidade, objetivos coletivos, interesses semelhantes e uma motivação intrínseca para “fazer bem”. De acordo com esta perspetiva, exigir às organizações que relatem resultados e outras informações de desempenho é apenas um mecanismo de intercâmbio que ajuda a melhorar o desempenho e a eficiência globais do setor. Segundo a teoria da *stewardship*, as informações sobre o desempenho ajudam a criar confiança e a garantir o alinhamento das metas de entidades financiadoras e organizações do terceiro setor. Isto, por sua vez, permite às primeiras alimentar relações

de longo prazo, com elevado desempenho e confiáveis com as segundas (Davis, Schoorman e Donaldson, 1997; Puyvelde *et al.*, 2012).

A perspectiva da **dependência de recursos**, por seu turno, assume que as organizações procuram os recursos necessários para a sua sobrevivência e por isso terão que estimular interações e trocas nos meios onde existem ou são controlados os recursos necessários (Pfeffer e Salancik, 2003). Os mecanismos de avaliação são usados não só por financiadores para controlar o desempenho das OTS, mas também por OTS para aceder a recursos, divulgando as suas atividades e respetivos resultados. As relações de dependência de recursos são, pois, vistas como interdependentes e mútuas: há uma interdependência de recursos (embora muitas vezes assimétrica), no qual as OTS dependem de doadores para aceder a recursos financeiros, utilizando a avaliação enquanto forma de dominar os recursos do meio, e doadores dependem de OTS para alimentarem a sua reputação e ganharem legitimidade (Ebrahim, 2002). A motivação das pessoas que dirigem as organizações será assegurar a sua sobrevivência, mantendo ao mesmo tempo a estabilidade nas suas relações de troca. Contrariando o pressuposto da teoria da agência de que as OTS são reativas em relação às expectativas externas de prestação de contas, a teoria da dependência de recursos enfatiza as estratégias proactivas que podem ser mobilizadas para lidar com o contexto em que operam (Faffee, 2001 *apud* Bies, 2010: 1067). No contexto desta perspectiva, a motivação para a criação ou participação em mecanismos de *accountability* vem das próprias organizações do terceiro setor. O empenho das organizações em controlar favoravelmente os recursos do meio fomenta o seu espírito empreendedor no campo da responsabilização enquanto via para aceder e preservar recursos (Bies, 2010).

Uma outra teoria com potencial explicativo para a presença de um sistema de prestação de contas é a **teoria institucional**, de acordo com a qual, o meio molda a estrutura e comportamento organizacional. DiMaggio e Powell (1983) falam desta pressão institucional, distinguindo entre isomorfismo coercivo – produto das relações de poder e dependência que surgem dos processos de contratação com o Estado – isomorfismo mimético – manifesto quando as organizações se modelam umas às outras na resposta à incerteza (simbólica), e isomorfismo normativo – relacionado com os processos de

profissionalização. As pressões normativas, miméticas e coercivas do ambiente compelem as organizações a conformar-se aos padrões e crenças convencionais, que são reforçados por forças institucionais locais (DiMaggio e Powell, 1983). Cada ambiente institucional conta com pressões, regras, normas, requisitos e sanções a que as organizações individuais têm que se conformar para receberem apoio e legitimidade. A perspectiva institucional postula que a compreensão do comportamento de uma organização requer a compreensão do seu ambiente externo, oferecendo uma lente para as dimensões estruturais do modelo de prestação de contas *vis-à-vis* a influência das instituições externas nas abordagens da *accountability* que as OTS escolhem. As organizações poderão pois alcançar legitimidade ajustando-se estruturalmente às influências institucionais. Para o efeito, podem utilizar diversos métodos, como seja a obediência às diretrizes institucionais, a reprodução das respostas de outras organizações a estas influências, ou a conformação com as normas e rituais institucionais (DiMaggio e Powell, 1983). Embora possam ser usados diferentes métodos, todos eles traduzem uma reação necessária às influências externas. A não conformidade com as expectativas institucionais relativamente à prestação de contas pode afetar negativamente a legitimidade organizacional (Bies, 2010).

2.5.2. Quadros interpretativos da adesão e uso da avaliação: entre influências externas e internas

A breve apresentação anterior deixa transparecer que a *accountability* não constitui uma panaceia social simples e clara, mas antes um constructo complexo e de alguma forma ambíguo (Ebrahim, 2005: 60). Não sendo os únicos, os esforços e mecanismos de avaliação são um reflexo das relações entre atores organizacionais arreigados num ambiente social e institucional.

Embora os pressupostos das teorias acima explicitadas não sejam mutuamente exclusivos, cada uma delas oferece uma contribuição única para explicar porque é que as organizações fazem avaliação e como usam os seus resultados. Os pressupostos diferem, no entanto, sobretudo de duas maneiras: a primeira tem a ver com os **fatores explicativos** da adoção de práticas de avaliação (ou, dito de outro modo, motivação para avaliar), e pode desdobrar-se em dois eixos – a pressão externa para avaliar (Carman, 2011; Højlund,

2014) e a propensão interna para fazer avaliação (Højlund, 2014) –, e a segunda tem a ver com a forma como a **informação de avaliação é utilizada** (Carman, 2011; Højlund, 2014).

A ideia de que as pressões para desenvolvimento de processos de avaliação nas organizações do terceiro setor decorrem de **forças externas e internas** encontra-se muito presente na literatura (Lewis, 2001:135-136; Hudson, 2003: 70; Perrin, 2006: 21-22; Carman e Fredericks, 2008: 52; Nigel e Smith, 2010: 20-21). Em primeiro lugar, ressalta na literatura a ideia da crescente pressão das entidades financiadoras/tutelares - como sejam o Estado, organismos/instituições da União Europeia, ou outras organizações Internacionais – e outros *stakeholders*, como sejam clientes, outras organizações sociais e a comunidade como um todo – no sentido de as OTS demonstrarem a sua efetividade e documentarem os resultados das intervenções apoiadas. As exigências de transparência e *accountability* são crescentes, e as entidades financiadoras querem saber o que é que foi realizado com os recursos concedidos, além de que existe uma maior disposição para assumirem os custos associados à avaliação. Em segundo lugar, é aludido um conjunto de pressões que advêm do interior das próprias organizações. Desde logo, reconhece-se que a avaliação é do interesse das OTS, uma vez que os ensinamentos retirados podem ser incorporados no processo de decisão, no sentido de melhorar o desempenho futuro, e também porque, mostrando que esses recursos são bem empregues, a confiança pública é reforçada e os apoios poderão ser intensificados, ainda que não seja claro que a avaliação conduza efetivamente a um aumento do financiamento às OTS (Harlock, 2013).

Partindo do enquadramento proposto por Carman (2011) para a compreensão das diferentes experiências das OTS com a avaliação, associam-se as teorias da dependência e da *stewardship* à **motivação interna** para fazer avaliação (motivação que vem do interior da organização – da direção ou gestão), enquanto as teorias institucional e da agência se associam sobretudo à **motivação externa** (motivação que vem do exterior da organização – entidades financiadoras ou outras filiações institucionais). A teoria dos *stakeholders* (não incluída no modelo de Carman), por seu turno, ao integrar elementos das restantes quatro teorias, poderá associar-se quer à motivação interna quer à motivação externa para fazer avaliação. Em duas destas teorias o uso da avaliação vincula-se à **alavancagem/promoção (*leverage*)** ou à **obtenção de algo em troca**, como seja ganhar legitimidade (teoria

institucional) ou assegurar recursos e reduzir a incerteza (teoria da dependência). Para duas das outras teorias, o uso está relacionado com a **prestação de contas** – satisfazer os requisitos das entidades financiadoras (teoria da agência) e asseverar a confiança e a qualidade do serviço (teoria da administração). Por fim, para a teoria das partes interessadas o uso tanto pode estar associado à prestação de contas (satisfazer exigências dos *stakeholders*, como à obtenção de algo em troca (legitimidade, controlo sobre os recursos) (vide quadro 8).

Quadro 8. Contributos teóricos para a compreensão das motivações e usos da avaliação

Uso	Motivação	
Alavancagem	<p>Externa</p> <p>Teoria Institucional</p> <p>Premissa: Pressões normativas, miméticas ou coercivas do meio ambiente moldam as operações internas das OTS.</p> <p>Objetivo: Obter legitimidade de atores externos.</p>	<p>Interna</p> <p>Teoria da dependência</p> <p>Premissa: As organizações têm que conseguir recursos de fontes externas para sobreviverem.</p> <p>Objetivo: Reduzir a incerteza e assegurar recursos.</p>
	<p>Teoria das partes interessadas</p> <p>Premissa: As OTS têm múltiplos stakeholders e são “responsáveis” de formas complexas e diversas perante uma variedade de grupos e interesses.</p> <p>Objetivo: Satisfazer os desígnios dos constituintes; obter legitimidade; assegurar a sobrevivência.</p>	
Prestação de contas	<p>Teoria da agência</p> <p>Premissa: Baseia-se nas relações de controlo e supervisão, por vezes adversárias, entre <i>principal</i> (ent. financiadoras) e <i>agente</i> (OTS).</p> <p>Objetivo: Satisfazer as necessidades de informação das entidades financiadoras.</p>	<p>Teoria da <i>stewardship</i></p> <p>Premissa: Relações entre principal e agente podem ser mutuamente benéficas.</p> <p>Objetivo: Assegurar confiança e qualidade do serviço.</p>

Fonte: Adaptado de Carman, Joanne G. (2011) “Understanding Evaluation in Nonprofit Organizations” *Public Performance & Management Review*. 34 (3), 350–77, Quadro 1.

Apoiando-se na teoria institucional, Højlund (2014) vem complexificar e matizar estas relações, entendendo a **pressão externa** e a **propensão interna para avaliar** enquanto fatores relacionados, mas autónomos⁵⁰. Segundo o autor, a adoção de práticas de avaliação não depende apenas da pressão e das condições do ambiente mas também de **fatores internos da organização**. Retomando a análise de Boswell (2008), Højlund usa a

⁵⁰ No enquadramento anterior a motivação interna e externa constituem categoria de uma mesma variável – a motivação.

distinção de Brunsson (2002, *apud* Højlund, 2014) entre "organização de ação" e "organização política", para rotular os dois extremos no *continuum* da propensão interna para avaliar. A "organização de ação", que poderemos associar à organização prestadora de serviços, que faz assentar a sua legitimidade na capacidade de produzir *outputs*, é associada a uma elevada propensão interna para avaliar, enquanto a "organização política", que podemos associar à organização de advocacia ou expressão de interesses, é associada à baixa propensão interna para avaliar uma vez que, fazendo assentar a sua legitimidade nas decisões políticas e na tomada de posições sobre os assuntos, gera mudanças de difícil mensuração e, portanto, mais difíceis de avaliar. Enquanto o primeiro tipo de organização tenderá a fazer um uso instrumental do conhecimento para melhorar a sua eficiência e assim aumentar a sua legitimidade, as organizações políticas tendem a usar o conhecimento simbolicamente e, não instrumentalmente, ao legitimar-se (Højlund, 2014: 34). Estes dois tipos de organizações e a ideia da sua associação a práticas e usos distintos da avaliação encontram algum suporte nas conclusões do estudo de OTS conduzido por Eckert e Moulton (2011), segundo o qual as organizações prestadoras de serviços se associam particularmente ao uso instrumental dos resultados da avaliação, enquanto as organizações de advocacia política, cujos efeitos são particularmente difíceis de medir, se associam especialmente aos usos simbólicos.

Højlund (2014) faz corresponder os **diferentes processos de adaptação institucional** (a que DiMaggio e Powell se referem como isomorfismo) a diferentes graus de pressão externa e a **diferentes níveis de propensão interna para avaliar**, sendo que um elevado grau de pressão externa para fazer avaliação não corresponde necessariamente a uma baixa propensão interna para avaliar. Com efeito, aos processos de adaptação institucional mimético e normativo associam-se, respetivamente, elevados e baixos níveis quer de pressão externa quer de propensão interna para avaliar, enquanto na adoção coerciva – e na adoção voluntária⁵¹ – se verifica uma relação inversa entre ambos os fatores: à maior pressão externa corresponde uma menor propensão interna para avaliar, no primeiro caso, e à menor pressão externa corresponde uma maior propensão interna

⁵¹ O autor inclui também no modelo a "adoção voluntária" que representa a "pura organização orientada para a ação" que, voluntariamente, adota práticas de avaliação quando necessário para aumentar a eficiência e os *outputs*, muito em linha com a ideia "comum" sobre avaliação (Højlund, 2014: 35).

para avaliar, no segundo caso. O autor faz ainda corresponder aos **quatro tipos ideais de modos de adoção da avaliação** aos tipos de **usos** esperados para cada um deles. Com efeito, aos modos de adaptação coercivo e normativo (a que correspondem níveis baixos de propensão interna para avaliar) o autor associa o não uso ou usos não instrumentais (simbólico e concetual, respetivamente), enquanto aos modos de adoção mimético e voluntário (a que correspondem níveis elevados de propensão interna para avaliar) estão associados usos instrumentais, respetivamente, de legitimação e concetual.

Embora a adoção de práticas de avaliação possa ser primeiramente influenciada por fatores externos, uma OTS estratégica quererá maximizar os seus recursos, procurando beneficiar do uso da avaliação no sentido de promover a eficácia e assegurar a sua sobrevivência. Esta expectativa está em linha com **a teoria da contingência**⁵², segundo a qual a eficácia pode ser vista como o produto do ajustamento entre a estrutura e papéis internos e o ambiente externo das organizações (Lawrence e Lorsch, 1967 *apud* Eckert e Moulton, 2011; Herman e Renz, 2008). Neste enquadramento, as organizações poderão usar a avaliação para se tornarem mais eficazes, mediante a melhoria do alinhamento entre o ambiente externo e interno (Eckert e Moulton, 2011). Donde, a compreensão da diversidade de práticas e usos da avaliação nas OTS poderá ser informada não só pelas suas principais funções e papéis⁵³, mas também por uma série de outros fatores contingenciais, nomeadamente a dimensão e a antiguidade. A teoria da contingência permitirá, assim, olhar para algumas dimensões internas associadas à adoção e ao uso das avaliações e para as suas interações.

Apesar de classificadas sob o mesmo setor, as OTS constituem um grupo heterogéneo, difuso e fragmentado, representando propósitos, missões, valores diversos e desempenhando papéis distintos na sociedade. Marshall (1996) sugere mesmo que o TS

⁵² A teoria da contingência considera as organizações enquanto sistemas de partes inter-relacionadas, enfatiza a importância dos fatores ambientais e sugere que não existe uma única forma de gerir melhor. Contrastando com a gestão científica, a teoria da contingência sustenta que, em vez de se procurarem princípios universais que se apliquem a todas ou à maioria das organizações, devem identificar-se os princípios de contingência que refletem as exigências de tipos específicos de “task environments” em que atuam. O termo “task environments” refere-se aos elementos específicos com que as organizações interagem no curso das suas operações (Anheier, 2005: 148).

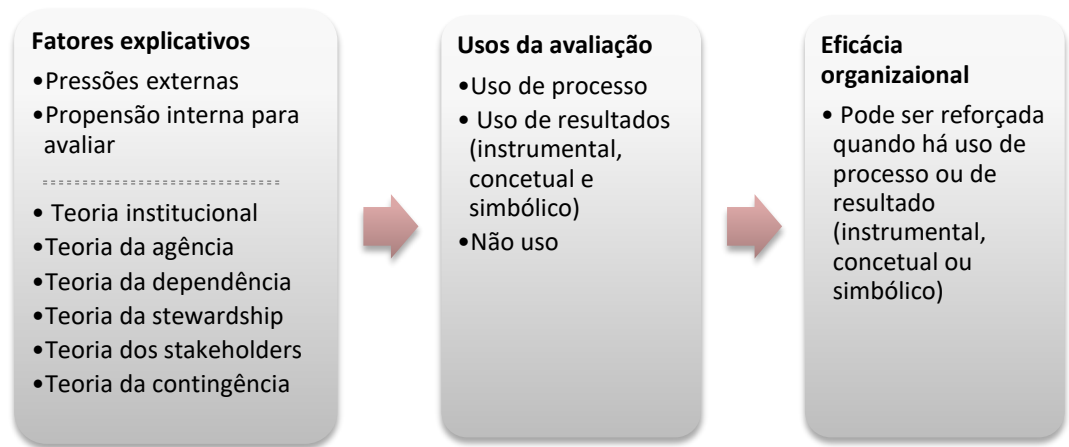
⁵³ O “task environment” das organizações releva, em primeiro lugar, a natureza do produto ou serviço prestado (Anheier, 2005: 152).

é demasiado heterogéneo para ser identificado a partir de qualquer conjunto específico de características, e conclui que a heterogeneidade interna do setor apenas pode ser abarcada se este for visto como uma amálgama de vários subsectores, cada um com um ímpeto e um carácter diferente. Em todo o caso, o seu carácter heterogéneo torna a generalização e a uniformização desafiante. É difícil comparar, por exemplo, uma associação recreativa, cultural e desportiva com uma casa abrigo para vítimas de violência doméstica. Todavia, agrupando-as no mesmo setor, a equivalência é frequentemente assumida e as práticas de avaliação são amiúde prescritas uniformemente. É por isso que ao invés de uma abordagem “one-size-fits-all”, são muitas as análises que defendem a pertinência e adequação de abordagens diferenciadas e multidimensionais (por exemplo, Herman e Renz, 1999; Eckerd e Moulton, 2010; Breckell, Harrison e Robert, 2010).

Ebrahim e Rangan (2010) rejeitam as abordagens de avaliação uniformizadas e propõem um enquadramento contingencial para a avaliação do desempenho social das organizações. A essência do modelo para a avaliação do desempenho social das organizações proposto por Ebrahim e Rangan (2010) é exatamente a contingência, sugerindo que dada a diversidade das atividades, objetivos e capacidades das organizações do setor social, e dada a complexidade da teoria da mudança e estratégia operacional de uma organização, para algumas organizações faz sentido medir as mudanças de longo prazo enquanto para outras faz mais sentido ficarem-se pela mensuração dos resultados de curto prazo.

Partindo dos pressupostos de **uma diversidade de teorias organizacionais** e inspirada pelos modelos interpretativos da adoção e uso da avaliação propostos por Carman (2011) e por Højlund (2014), é possível estabelecer o quadro das ligações que informam a eficácia organizacional (vide figura 3), e que orientam a minha análise das motivações, propósitos e usos nas OTS portuguesas.

Figura 3. Relações entre fatores explicativos da avaliação, uso da avaliação e eficácia organizacional



Fonte: Elaboração da autora.

Partindo das críticas formuladas a uma diversidade de abordagens e debates atuais sobre a adoção e usos da avaliação, e seguindo de perto a proposta, **adoto neste trabalho uma perspectiva ampla e integrada sobre avaliação, enquanto mecanismo de prestação de contas**. A perspectiva que presidirá a esta pesquisa é a de que as OTS são mais complexas do que as perspectivas convencionais da relação *principal/agente* sugerem, uma vez que a **ação das OTS é influenciada não só por múltiplos principais** e interesses, mas também pelos seus próprios **valores e missão**. Seguindo esta linha, “as missões adicionarão uma dimensão ética e ancorada em valores à responsabilização uma vez que enfatizam as motivações internas dos atores em detrimento das pressões externas exercidas pelos *principais*” (Ebrahim, 2003b: 199). A **prestação de contas** será, nesta perspectiva, tanto **interna** como **externa**, envolvendo não apenas controlo externo, mas também a “responsabilidade sentida” (Fry, 2005). Como tal, espera-se que não seja apenas uma resposta reativa à supervisão, de reação ao que é imposto, mas também uma resposta proactiva, ligada ao cumprimento da missão e à garantia de que a confiança pública é servida.

A relevância do enfoque na dimensão interna da avaliação é também expressa pela **heterogeneidade de funções e papéis das OTS**, cuja eficácia deverá ser vista como o produto do **ajustamento entre a estrutura e papéis internos e o ambiente externo das organizações** e não apenas como reflexo das pressões institucionais.

Sublinha-se, em todo o caso, a **natureza relacional da avaliação** (Ebrahim, 2003a, 2003b; Bies, 2010), que poderá apenas ser entendida no contexto das interações das entidades do terceiro setor com os membros do seu ambiente organizacional, como sejam financiadores/as, parceiros, beneficiários/as e comunidade. Se a prestação de contas está ancorada nas relações entre atores organizacionais, a avaliação não pode ser compreendida adequadamente sem alguma consideração sobre para quem e para quê eles são empregues. A prestação de contas a financiadores/as, operacionalizada através de relatórios ou outras formas de controlo pode estar em desalinho com a responsabilização perante a missão. A avaliação pode ser vista, neste sentido, como um “produto de relações intra e interorganizacionais e tendencialmente enviesada em favor dos atores dominantes” (Ebrahim, 2003b: 199).

Esta aceção vem reafirmar o carácter político da avaliação sobre avaliação e a importância de ter em conta a sua **imersão nas relações de poder** (Greene, 2006; Chapman *et. al*, 2010; Arvidson e Lyon, 2014). No entender de Ebrahim (2005: 82), a discussão política sobre como reforçar a responsabilização através da avaliação ou medição do desempenho/impacto pode ser equivocada se ela assentar e privilegiar os mecanismos e atores ascendentes em detrimento dos descendentes e internos. Tal não significa que a avaliação para prestação de contas ascendente é desnecessária, uma vez que **desempenha um papel crucial e legítimo**, por exemplo, na prevenção da utilização fraudulenta de fundos por organizações, mas é apenas uma dimensão das múltiplas relações de prestação de contas.

Depois, e em termos relativos, a avaliação deve ser vista mais como enquadrando um sistema de **relações multidirecionais e contingentes** do que como um conjunto de ligações binárias independentes. Isso significa que o reforço dos mecanismos de análise do desempenho não respeitam apenas à prestação de contas a financiadores, mas também ao avançar no prosseguimento da missão, refletindo a prestação de contas perante comunidades e clientes. É, pois, importante estudar os diversos mecanismos de responsabilização ascendente, descendente e interna, particularmente a sua natureza contingente e a forma como variam de acordo com o propósito, direção e contexto.

Um desafio central para as OTS é encontrar um **equilíbrio** entre os mecanismos que respondem aos desígnios de **responsabilização ascendente** das entidades financiadoras e aos que respondem às necessidades dos membros e comunidade (i.e., **responsabilização interna e descendente**), enquanto desencadeiam também mudanças positivas no comportamento organizacional. Para muitas organizações, encontrar este equilíbrio exigirá uma reorientação para processos de aprendizagem e responsabilização perante a missão no sentido de colocar a avaliação de sentido ascendente em perspectiva adequada, assim como a conjugação de processos de avaliação de curto e longo prazo. A questão chave é que a prestação de contas, tal com o poder, é um conceito relacional, pelo que os efeitos dos seus mecanismos apenas podem ser compreendidos quando colocados em contexto. Nos casos em que as entidades financiadoras partilham com as OTS uma perspectiva de longo prazo, este equilíbrio poderá ocorrer com menor dificuldade. Já nos casos em que as prioridades daquelas enfatizam os resultados de curto prazo à custa da aprendizagem de longo prazo, podem ter lugar prestações de contas conflitantes (Ebrahim, 2005: 73).

Ainda que avaliação seja frequentemente discutida como sendo imposta às OTS do exterior e apesar dos seus enviesamentos, uma prestação de contas com base no desempenho forte é essencial para qualquer organização do terceiro setor. É necessária para a sua própria sobrevivência porque, tal como sinalizam Edwards e Hulme (1995:6), **“trabalhar com efetividade e prestar contas com transparência, são componentes essenciais de uma prática responsável**, de que depende em última análise a legitimidade da intervenção para o desenvolvimento”.

2.6. Cultura de avaliação e aprendizagem organizacional

A forma como e a extensão com que os resultados da avaliação são utilizados têm constituído uma grande preocupação para quem faz e para quem consome a avaliação. As primeiras avaliações, após a segunda guerra mundial (sobretudo nos EUA), particularmente estudos encomendados de programas educacionais de larga escala, resultavam frequentemente na sua não utilização. O incómodo relativamente a esta situação incitou a investigação sobre os vários tipos de uso da avaliação (i.e., instrumental, concetual,

política) e os fatores que os influenciam (p.ex., oportunidade, relevância, características de quem faz as avaliações e de quem as utiliza) (Patton, 1978; Weiss, 1981, entre outros). Por outro lado, os estudos no campo da avaliação, particularmente as abordagens enquadradas no **paradigma pragmático**, passaram a dedicar atenção à relação entre quem avalia e quem usa a avaliação. Um produto deste trabalho tem sido o desenvolvimento das **abordagens participativas/colaborativas à avaliação** (por exemplo, Cousins e Earl, 1992; Patton, 1997, 2013; Cousins e Whitmore, 1998; Cousins, 2013). Estas abordagens reivindicam, tanto quanto possível, a participação dos *stakeholders* no desenho e recolha de dados da avaliação. O envolvimento de *stakeholders* no desenho e implementação da avaliação destina-se a aumentar: (a) a sua adesão à avaliação; (b) a sua compreensão do processo de avaliação; e (c) potenciar a sua utilização dos resultados da avaliação (Torres e Preskill, 2001: 388). Michael Quinn Patton (1997) classifica a aprendizagem que decorre desta participação de “uso do processo”.

Apesar de as abordagens participativas terem vindo a ser amplamente associadas ao aumento da relevância e utilização dos processos e resultados da avaliação de um conjunto de programas e organizações, a participação dos utilizadores no desenho e atividades de avaliação é necessária mas não suficiente para realizar todo o potencial da avaliação no seio das organizações (Torres e Preskill, 2001: 388). Nos últimos anos, as equipas de avaliação têm-se concentrado na compreensão do papel e da prática de avaliação no processo de **aprendizagem das organizações**, como um todo (por exemplo, Lambert e Owen, 1995; 2013; Preskill e Torres, 1999; Torres e Preskill, 2001; Botcheva, White e Huffman, 2002; Ebrahim, 2005; Rogers e Williams, 2006; Hoole e Patterson, 2008).

Aprendizagem é essencialmente a forma através da qual os indivíduos atribuem significado a partir da informação que encontram. A aprendizagem organizacional constitui, nesta linha, um **conjunto de procedimentos e perspetivas que permite que os ensinamentos sejam efetivamente transferida para o ambiente organizacional** ou programa (Rogers e Williams, 2006: 77). Numa perspetiva complementar, Preskill e Torres (1999) definem aprendizagem organizacional como um processo contínuo de crescimento e melhoria que utiliza informação ou *feedback* acerca de processos e resultados para introduzir mudanças; está integrado nas atividades de trabalho e na infraestrutura

organizacional (na sua cultura, liderança, sistemas, estruturas, e mecanismos de comunicação) e invoca o alinhamento de valores, atitudes e percepções entre membros das organizações.

Conforme propõe Ebrahim (2005), a relação entre avaliação e aprendizagem é simultaneamente prática e normativa: no sentido de as avaliações serem úteis para as organizações que estão a ser avaliadas, **sendo necessário encontrar formas sistemáticas de garantir que a informação reverta para a tomada de decisão**. A aprendizagem, como tal, envolve a geração de conhecimento através do processamento de informação que depois será utilizada para mudar comportamentos (Ebrahim, 2005: 67). De acordo com esta utilização da aprendizagem, a simples geração de informação não é suficiente, devendo envolver também o uso do conhecimento no sentido de influenciar as práticas organizacionais. Estas observações reivindicam a necessidade de identificar e analisar **meios de interligação** dos mecanismos de geração de conhecimento, como seja a avaliação, com a aprendizagem, explicitamente.

Dentre os fatores que têm vindo a ser apontados como facilitadores da avaliação, sobressai o **papel da cultura organizacional**, a qual constitui tanto uma faceta do ambiente em que a organização atua como uma base para uma possível mudança. **A cultura organizacional compreende valores, ideias, crenças, pressupostos, normas, artefatos e padrões de comportamento compartilhados** (Schein, 1996)⁵⁴. Apesar de, como realçam Rogers e Williams (2006), a descrição plena de uma cultura organizacional que suporta a avaliação poder ser um ideal utópico, a extensão em que a efetiva cultura organizacional difere deste ideal pode ser útil para identificar o âmbito da avaliação que pode realmente contribuir para a melhoria das práticas, assim como para identificar os aspetos culturais inerentes à construção da capacidade avaliativa.

Para Preskill e Torres (1999), **a avaliação é uma componente central de uma organização aberta à aprendizagem** e o seu sucesso depende grandemente da cultura das organizações. Uma cultura de aprendizagem promove a assunção de risco, a aprendizagem a partir dos erros e um clima de confiança e encorajamento. Ao contrário, uma “cultura

⁵⁴. O mesmo autor identifica **três níveis** na cultura organizacional: **artefatos manifestos** (padrões de comportamento visíveis), **valores reportados** (estratégias, objetivos e filosofias conscientes, formais ou informais) e **valores básicos** (orientações tácitas de como uma organização funciona).

anti-aprendizagem” caracteriza-se pela resistência por parte do pessoal, receio de assumir riscos, atitudes negativas perante a recolha de dados e falta de confiança na organização.

Os principais fatores facilitadores ou obstaculizadores do desenvolvimento de uma cultura de avaliação consistente têm vindo a ser sintetizados em diversas análises. John Mayne (2010: 6-7) defende que uma cultura avaliativa forte exige um compromisso sério com processos de autorreflexão e autoanálise, bem como uma aprendizagem contínua baseada em evidências. O autor vincula a cultura de avaliação a uma cultura organizacional que procura deliberadamente informações empíricas sobre seu desempenho, a fim de utilizar essa informação para aprender a gerir melhor seus programas e serviços e assim melhorar seu desempenho.

Segundo Torres e Preskill (2001), por seu turno, **a aprendizagem organizacional envolve:**

- Estabelecimento de equilíbrio entre os papéis de *accountability* e aprendizagem da avaliação;
- Integração da função da avaliação na organização;
- Desenvolvimento de modelos de trabalho que relacionem resultados sobre programas e iniciativas específicos com metas globais das organizações;
- Sustentação de um espírito de diagnóstico contínuo que reclame uma aprendizagem incremental e iterativa/reiterada ao longo do tempo;
- Disponibilização de tempo para a reflexão, análise dos pressupostos tácitos, e diálogo entre quem avalia, intervenientes nos projetos e líderes organizacionais;
- Reconsideração dos papéis tradicionais de quem avalia e das competências de que necessitam.

No entender das autoras, quando estas circunstâncias estão reunidas, as organizações têm a oportunidade de implementar cabalmente a avaliação ao serviço das suas missões.

No enquadramento destas perspetivas, a **cultura de avaliação** numa organização está diretamente relacionada não só com as **práticas** de avaliação ao nível organizacional, mas também com uma **atitude individual e coletiva** de compreensão, aceitação, promoção

e valorização da avaliação enquanto processo contínuo de contributo para a mudança no sentido de melhoria da eficácia das organizações.

Tal como sugere Ebrahim (2005: 71), associar aprendizagem à avaliação requer atenção explícita à forma como a informação gerada pelas avaliações reverte nos processos de decisão. Ademais, implica um olhar sobre a forma como as **relações de poder** influenciam o escrutínio das falhas, particularmente entre múltiplos constituintes. Finalmente, e mais genericamente, uma abordagem de aprendizagem à avaliação sugere a necessidade de uma mudança de perceção das avaliações enquanto cartões de verificação de desempenho para uma visão das avaliações enquanto recursos de melhoria.

Avaliadores/as e organizações com uma visão da avaliação orientada para a aprendizagem enfrentam, pois, **desafios** significativos, que incluem (Torres e Preskill, 2001: 389):

- Financiamento sedento de prestação de contas e legislação que impõe resultados em períodos de tempo curtos;
- Colaboradores e colaboradoras das organizações a quem é dado pouco tempo ou apoio para se dedicarem à reflexão e diálogo que convide ao questionamento sobre os pressupostos, crenças e valores dos programas, políticas e práticas das respetivas organizações;
- Líderes com pouca experiência em fundamentar as decisões em dados, que não sabem como articular conclusões sistemáticas com outras formas de informação nos processos de decisão;
- Intervenientes nos projetos sobrecarregados que continuam a ver a avaliação como prescindível, acessória ou para a qual têm pouco tempo;
- Avaliadores/as que tenham interesse em, ou a capacidade de, implementar a avaliação como meio para aprendizagem e mudança organizacionais;
- Pouco suporte para funções de replaneamento e/ou para influenciar a cultura organizacional no sentido de sustentar a aprendizagem organizacional;
- Membros de organizações que percecionam a avaliação como ameaça, mantendo-se desconfortáveis com o diálogo de grupo desenhado para facilitar a aprendizagem a partir da avaliação;

- Quadros dos níveis intermédios que pretendem iniciar trabalhos de avaliação e encontram dificuldade em conseguir suporte dos níveis hierárquicos superiores.

Assim, uma **cultura avaliativa mais fraca** poderá, designadamente (Mayne, 2010: 6):

- Reunir informação sobre resultados, mas restringir o seu uso ao reporte externo;
- Reconhecer a necessidade de aprender, mas não disponibilizar o tempo ou ocasiões estruturadas para fazer;
- Ter dificuldade em submeter-se a mudanças;
- Afirmar que procura evidências, mas desencorajar o desafio e o questionamento do *status quo*;
- Falar sobre a importância de obter resultados, mas olhar com desconfiança a assunção de riscos ou valorizar o seguimento de regras e processos.

Estes desafios sugerem que o esforço de equilibrar prestação de contas e aprendizagem requer não apenas mudanças na estrutura e sistemas de informação das OTS mas também alterações nas perceções sobre aprendizagem e sobre as falhas, assim como o comprometimento com o fomento de relações interorganizacionais e intraorganizacionais que favoreçam a pesquisa, a experimentação e a reflexão crítica. A maior valia de tal integração não reside simplesmente na melhoria das práticas existentes, mas na capacitação das OTS para alcançar as suas missões. Os processos conscientes de aprendizagem podem ajudar as OTS a avaliar criticamente e a orientar as suas estratégias de desenvolvimento social, assim como a ampliar a sua capacidade de influenciar as políticas e práticas públicas.

Capítulo 3. Estratégia metodológica: um plano misto de pesquisa

O propósito da revisão teórico-analítica conduzida nos capítulos anteriores foi sobretudo retirar de cada abordagem conceitos e ferramentas que melhor servissem para ler ou tornar inteligíveis as realidades sociais que se encontraram no trabalho empírico. Não sendo possível dar conta de todos os estudos e abordagens que têm procurado compreender e enquadrar a adoção de práticas de avaliação nas organizações, procurou-se expor algumas discussões relevantes que, direta ou indiretamente, influenciam as perspectivas que adoto neste trabalho. Entendo estas abordagens como um corpo de conhecimento que, não sendo uniforme, acaba por ser complementar, permitindo uma leitura mais rica e abrangente das realidades que se pretendem observar. O estudo da avaliação nas organizações do terceiro setor constitui um desígnio abrangente, mobilizador de múltiplos olhares. Reconhecer esta multidisciplinaridade e as diferentes contribuições de cada modelo analítico serve apenas para enriquecer a construção da problemática.

A moldura teórica plural que tem vindo a ser construída desde os capítulos iniciais deste trabalho serviu, sobretudo, para me ajudar a definir o objeto de pesquisa. Neste capítulo explicitam-se as opções metodológicas do estudo: apresenta-se o modelo de análise, no qual se articulam de forma operacional os marcos e as pistas retidos para orientar o trabalho de observação e de análise, e descrevem-se as fontes e estratégias de recolha e análise da informação.

3.1. Objetivos

Em Portugal, só muito recentemente se começou a discutir a relevância da adoção de práticas de avaliação nas OTS, designadamente em virtude das transformações ocorridas nas próprias organizações, tendo em vista os papéis que têm vindo a desempenhar, e dos efeitos da sua relação com as entidades financiadoras/tutelares.

Partindo do reconhecimento da crescente relevância do fenómeno, ainda escassamente estudado em Portugal, e da consciência das tensões que o atravessam, o presente estudo propõe-se, em termos gerais:

- Conhecer os padrões de avaliação adotados pelas organizações do terceiro setor (OTS) e determinar os fatores que têm influenciado a respetiva evolução.

Em termos específicos pretende-se:

- Contextualizar a extensão e o significado atribuído à avaliação na literatura sobre teoria de avaliação, teoria das organizações, gestão estratégica, estudos sobre terceiro setor, estudos sobre eficácia, etc.;
- Identificar as principais influências no sentido da profissionalização da gestão das OTS;
- Identificar práticas e discursos institucionais de pressão sobre as OTS para adotarem práticas de avaliação;
- Caracterizar as práticas de avaliação conduzidas pelas OTS em termos de modelos e abordagens adotados, propósitos, recursos e destinatários/as;
- Analisar as relações estabelecidas entre as entidades financiadoras e OTS;
- Analisar as perceções, representações e atitudes de atores-chave (pessoal dirigente e pessoal técnico de OTS, entidades financiadoras, consultores/as de avaliação) relativamente à avaliação;
- Determinar os efeitos da avaliação realizada pelas OTS em termos de aprendizagem, capacitação e efetividade organizacionais;
- Aferir sobre o contributo das entidades financiadoras para a expansão de uma cultura de avaliação nas organizações, discutindo especificamente a importância do Fundo Social Europeu (FSE) para esse contributo;
- Identificar as principais barreiras e os principais fatores críticos de sucesso da avaliação;
- Determinar marcos da evolução das práticas de avaliação.

3.2. Dimensões analíticas

Do ponto de vista metodológico, a prossecução dos objetivos acima referenciados implicou o recurso a uma diversidade de contributos teóricos e analíticos, provenientes de diferentes campos disciplinares (desde a teoria da avaliação - de programas -, à teoria organizacional, da administração e da gestão), com destaque para modelos e tipologias de avaliação, de eficácia organizacional e de prestação de contas delineados por diversos/as autores/as e para as teorias explicativas da adoção e uso da avaliação.

Considerando o contributo dos requisitos do **financiamento do FSE**⁵⁵ para a crescente relevância dos processos de avaliação nas OTS, decidiu-se adotar dois dos programas de intervenção por ele cofinanciados como terreno de trabalho: o Programa Operacional da Iniciativa Comunitária EQUAL (PIC EQUAL) e o Programa Operacional Potencial Humano (POPH). A escolha destes programas prendeu-se com a amplitude e abrangência, no caso do POPH, e pelo particular investimento na promoção e organização de práticas de avaliação dos projetos (edição de manuais de apoio, realização de sessões de esclarecimento), no caso do PIC EQUAL.

O Programa Operacional Potencial Humanos (POPH)⁵⁶ é um programa *mainstream* (não experimental) particularmente vocacionado para intervenções na área social, que teve lugar no período de programação que se seguiu ao do PIC EQUAL. Em fase de implementação durante a pesquisa empírica conduzida neste trabalho, é também potencialmente relevante no que toca à integração de práticas de gestão e monitorização nas OTS, por se tratar de um programa com grande abrangência e amplitude em termos do volume de investimento e de entidades do terceiro setor abrangidas na qualidade de beneficiárias ou promotoras de projetos. Apoiou mais de 31.000 projetos, os quais embora

⁵⁵ O Fundo Social Europeu (FSE) é um dos dois Fundos Estruturais da União Europeia (o outro é o Fundo Europeu de Desenvolvimento Regional – FEDER), que se enquadra no contexto geral da política de coesão económica e social comunitária, tendo vindo a apoiar muitas das políticas nacionais de educação, formação, emprego e inclusão social. Ao longo de quatro ciclos de programação dos Fundos Estruturais (1988-93, 1994-1999, 2000-2006 e 2007-2013), um grande número de organizações que trabalham na área social desenvolveram projetos em diversas áreas ao abrigo de um conjunto de programas e medidas, **constituindo uma importante fonte de financiamento para muitas entidades e para algumas respostas na área social**. A nível nacional, o Quadro de Referência Estratégica Nacional (QREN), que vigorou entre 2007 e 2013, permitiu o investimento em território nacional de 9.615 M€ direcionado especificamente para a Agenda Temática Potencial Humano (FSE) (<http://www.qren.pt/np4/3341.html>). A Agenda Temática Valorização do Território (FEDER), por seu turno, permitiu apoiar a construção ou requalificação de 408 equipamentos de apoio social, sendo na sua maioria creches e lares de idosos (<http://www.qren.pt/np4/3341.html>).

⁵⁶ Este programa sucedeu ao Programa Operacional Emprego, Formação e Desenvolvimento Social (POEFDS) do anterior período de programação (2000-2006). O PO Potencial Humano “é um programa multiobjectivo cujas orientações comunitárias vão no sentido de estimular o potencial de crescimento sustentado da economia portuguesa, no quadro das seguintes prioridades: Superar o défice estrutural de qualificações da população portuguesa, consagrando o nível secundário como referencial mínimo de qualificação, para todos; Promover o conhecimento científico, a inovação e a modernização do tecido produtivo, alinhados com a prioridade de transformação do modelo produtivo português assente no reforço das atividades de maior valor acrescentado; Estimular a criação e a qualidade do emprego, destacando a promoção do empreendedorismo e os mecanismos de apoio à transição para a vida ativa; Promover a igualdade de oportunidades, através do desenvolvimento de estratégias integradas e de base territorial para a inserção social de pessoas vulneráveis a trajetórias de exclusão social.” (POPH, 2015: 7).

não sendo promovidos apenas por organizações do terceiro setor, envolveram um grande número destas⁵⁷.

O PIC EQUAL foi um programa financiado pelo Fundo Social Europeu, executado no período 2000-2010⁵⁸, de carácter experimental, com o “objetivo de desenvolver abordagens inovadoras para combater as discriminações no acesso e no mercado de trabalho que, depois de validadas, deveriam ser generalizadas” (Vale, 2010: 287). Por se tratar de uma iniciativa de carácter experimental, teve uma amplitude naturalmente inferior à do contemporâneo PO Emprego, Formação e Desenvolvimento Social (POEFDS) e sucessor POPH, quer em termos de volume de financiamento, quer em termos de projetos aprovados e número de entidades beneficiárias. Ainda assim, foram envolvidas 1111 organizações no programa, 900 das quais organizações do terceiro setor, as quais deram corpo a mais de 200 parceiras de desenvolvimento de projetos inovadores.

Sistematizam-se algumas das dimensões em que se organiza a pesquisa e que inspiraram as perspetivas aventadas nos capítulos anteriores:

- Caracterização das organizações em termos de forma legal/estatuto jurídico, antiguidade, dimensão, recursos, vocação e visão estratégica enquanto elementos distintivos da identidade organizacional.

- Caracterização das práticas de avaliação conduzidas pelas OTS (intervenientes, momento das intervenções sobre as quais incide o processo avaliativo, âmbito de incidência da avaliação, metodologias utilizadas, agentes, destinatários da avaliação e envolvimento dos *stakeholders*).

- Análise dos principais propósitos e motivações (internos/externos) associados à adoção de práticas de avaliação nas organizações.

- Análise dos tipos de usos da avaliação prevalecentes, ponderando o seu contributo para a aprendizagem organizacional.

⁵⁷ Nos relatórios de execução do POPH, não encontramos informação sobre o número total de entidades da economia social beneficiárias dos apoios concedidos no âmbito do POPH, mas de acordo com listagens com informação parcial (de entidades com projetos contratualizados aprovados, em execução em junho de 2012) o número de entidades do TS beneficiárias deverá ascender às 1.300.

⁵⁸ Apesar de a execução do programa se estender até 2010, a implementação do PIC EQUAL integrou o ciclo de programação dos fundos estruturais de 2000-2006, não tendo continuidade no ciclo de programação subsequente (e no qual teve enquadramento o POPH).

- Análise dos principais momentos e marcos associados ao surgimento, desenvolvimento e consolidação ou declínio das práticas de avaliação.
- Análise das representações e atitudes de atores-chave relativamente à avaliação (dirigentes e técnicos de OTS, entidades financiadoras, consultores/as de avaliação).
- Análise das relações entre entidades financiadoras e OTS (de reforço/ conflito/ independência) e entre OTS e consultores/as de avaliação.
- Análise das racionalidades inerentes aos objetivos e requisitos de avaliação do POPH e EQUAL assim como do funcionamento das respetivas unidades técnicas de gestão e acompanhamento, numa perspetiva comparada.
- Análise do contributo do POPH e PIC EQUAL para o desenvolvimento e sustentabilidade das práticas de avaliação nas OTS, numa perspetiva comparada.
- Análise dos constrangimentos e fatores críticos de sucesso dos processos de avaliação nas OTS.

Matriz do processo de produção de informação

EIXOS ANALÍTICOS	Conceitos	Dimensões	INDICADORES
Identidade organizacional	Dimensão Autonomia Vocação	Financeira Recursos Humanos Missão Papéis	- Receitas e fontes de receitas da organização - N.º pessoas ao serviço - Acesso a recursos financeiros necessários para a prossecução da estratégia definida - Formas e estatutos jurídicos - Antiguidade - Causas/objetivos, visão e estratégia da organização - Públicos-alvo - Setor/Perfil da atividade - Planos da organização, em termos de atividades
Processos e atividades de avaliação nas OTS	Recursos de avaliação Estruturas de avaliação	Financeiros Humanos Competências Materiais	- Tempo, recursos humanos e financeiros dedicados à avaliação - <i>Expertise</i> em avaliação - Ferramentas informáticas de suporte - Suporte técnico e financeiro de <i>stakeholders</i> externos - Departamento ou estrutura interna própria dedicada à avaliação

	Tipos de avaliação	<p>Abordagens e Metodologias de avaliação</p> <p>Níveis de avaliação</p> <p>Escala/dimensão dos objetos de avaliação</p> <p>Incidência/conteúdos da avaliação</p> <p>Agentes da avaliação</p> <p>Participantes na avaliação</p> <p>Destinatários/as da avaliação</p> <p>Fluxo e cobertura</p> <p>Propósitos da avaliação</p>	<ul style="list-style-type: none"> - Diretrizes e regras comuns para a realização de avaliações - Definição de funções e responsabilidades associadas à avaliação - Avaliação realizada de forma rotineira/espórádica - Práticas de gestão e avaliação: <ul style="list-style-type: none"> . Relatórios anuais de contas/atividades . Relatórios para entidades financiadoras sobre atividades de projetos . Relatórios para entidades financiadoras sobre execução financeira de projetos . Auditorias financeiras/técnicas . Certificação/acreditação . Avaliações de desempenho do pessoal ao serviço . Monitorização de projetos . Receção de visitas no local de financiadores e entidades reguladoras . Monitorização do cumprimento dos objetivos e metas dos projetos . Planeamento estratégico - Utilização de um sistema de gestão <i>"balanced scorecard"</i> . Avaliações formais de projetos/do desempenho organizacional . Conceção de "modelos lógicos" de programas - Recolha de Informação de monitorização/avaliação sobre: <ul style="list-style-type: none"> . Despesas das intervenções . Nº/perfil/satisfação de beneficiários/as . <i>Outputs/</i> efeitos/impactos das intervenções . Boas práticas, padrões de referência . Informação de controlo ou comparação . Narrativas ou testemunhos pontuais - Métodos empregues na produção de informação de avaliação - Principal responsável pela dinamização dos trabalhos de avaliação - Recurso a avaliação externa - Atores envolvidos nos processos de avaliação - Procedimentos de follow-up
--	--------------------	--	--

	Cultura organizacional	Recetividade Suporte institucional	<ul style="list-style-type: none"> - Procedimentos para articular os resultados da avaliação - Atores a quem os resultados da avaliação são apresentados - Uso dos resultados da avaliação - Perceção da avaliação pelos membros da organização (oportunidade de aprendizagem/ameaça/desperdício) - Suporte da liderança organizacional (- Perceção sobre requisitos de avaliação das entidades financiadoras (adequados/inadequados) - Atitude perante a produção de dados
Fatores explicativos da adoção de práticas de avaliação	Motivação	Interna Externa	<ul style="list-style-type: none"> - Razões para adoção de práticas de avaliação - Uso dos resultados da avaliação - Visões sobre avaliação
Impacto da avaliação nas OTS	Aprendizagem organizacional	Tomada de decisão Prestação de contas Mudança organizacional Pressupostos	<p>Utilização dos resultados da avaliação para:</p> <ul style="list-style-type: none"> - decidir sobre alocação de recursos (humanos, financeiros, materiais) - planejar/rever projetos/programas/estratégias - alterar/ajustar procedimentos organizacionais - fazer escolhas estratégicas sobre o futuro da organização <p>reportar entidades financiadoras/tutelares</p> <ul style="list-style-type: none"> - reportar aos órgãos sociais/direção - divulgação e relações públicas <p>Perceção do contributo dos resultados de avaliação para:</p> <ul style="list-style-type: none"> - qualidade dos serviços prestados - captação de novos financiamentos - reconhecimento da comunidade - satisfação dos utentes/participantes - inovação nos produtos e serviços - reorientação da estratégia da organização - ajudar a obter novos financiamentos - revisão de projetos - Institucionalização da avaliação como parte do sistema de informação organizacional, estrutura de poder, processos e sistemas que influenciam a tomada de decisão e a ação

utilitaristas) que, no primeiro capítulo, foram organizadas em função do paradigma científico em que se enquadram (positivista, construtivista, transformativo e pragmático). Além dos contributos da teoria da avaliação, alicerçados na pesquisa em ciências sociais, a produção e interpretação dos resultados do estudo nestas componentes foi feita à luz das perspetivas e modelos de gestão organizacional (NGP, gestão por objetivos, gestão orientada para resultados, gestão da qualidade, medição/gestão do desempenho, controlo, auditoria, ...), classificadas como abordagens managerialistas à avaliação, e das perspetivas e ferramentas de avaliação da eficácia/desempenho/impacto das OTS. A análise dos tipos de avaliação é inspirada pelos elementos das sistematizações de Guerra (2002), Capucha *et al.* (1996: 12-18) e Cuesta e Martínez (2010: 83), e pelos enquadramentos do processo de avaliação propostos por vários/as autores/as, designadamente Cutt e Murray (2000), Ebrahim e Rangan (2010, 1014) e Liket, Rey-Garcia e Maas (2014).

O modelo concebido para analisar a influência do contexto organizacional interno nas práticas de avaliação, em termos de recursos e estruturas organizacionais de suporte à avaliação e de cultura organizacional, é particularmente tributário dos modelos explicativos que descrevem e ilustram as condições necessárias para alicerçar a avaliação nas organizações (Stufflebeam, 2002; Valkov e King, 2005; Bourgeois e Cousins, 2008; Powell e Boyd, 2008; Preskill e Boyle, 2008; Bourgeois e Cousins, 2013), com destaque para abordagens que se centram na cultura e na aprendizagem organizacional, em especial, Preskill e Torres (1999), Cousins *et al.* (2004) e de Ebrahim (2005a).

A compreensão da adoção de práticas de avaliação nas OTS, seja das suas motivações internas, seja das associadas ao ambiente externo, ancora-se sobretudo nas abordagens económicas e organizacionais, nomeadamente das propostas da teoria da escolha racional, da teoria da agência, da perspetiva da dependência de recursos, das perspetivas das partes interessadas, da teoria da *stewardship* e da teoria institucional. Os trabalhos que exploram a influência destas teorias na adoção de práticas de avaliação nas organizações, em particular os conduzidos por Eckerd e Moulton (2010) Højlund (2014) e Carman (2011) são especialmente inspiradoras desta análise. Por seu turno, as discussões sobre os processos de *accountability* – designadamente as que focam na sua articulação

com a aprendizagem organizacional –, sobretudo as protagonizadas por Ebrahim (2003, 2005, 2007, 2009), são particularmente mobilizadas para a interpretação das relações entre entidades financiadoras (e outros *stakeholders*) e OTS.

Os usos da avaliação são analisados a partir das tipologias de uso delineadas em diversas análises, com destaque para o trabalho de Alkin e Taut (2003), no qual é abordada a distinção entre usos instrumentais e usos não instrumentais (conceituais e simbólicos/legitimadores) e de Patton (2011, 1997), que destaca ainda o uso de processo. As abordagens participativas/colaborativas à avaliação (por exemplo, Green, 1988; Cousins e Earl, 1992; Patton, 1997, 2013; Cousins e Whitmore, 1998; Cousins, 2013) são, por seu turno, inspiradoras da análise da influência do envolvimento dos *stakeholders* nos usos da avaliação.

As reflexões conduzidas nos diversos eixos e dimensões de análise são articuladas com as perspectivas teóricas sobre a existência, os papéis, as atividades e os contributos do terceiro setor, e com as discussões sobre a sua identidade e especificidades enquanto setor distinto da economia de mercado e do Estado, presentes na vasta literatura internacional e nacional sobre o setor (por exemplo, Lewis, 2001; Anheier, 2005; Evers e Laville, 2004; Edwards e Folwer, 2002; Ferreira, 2008, 2009, 2012, 2015).

Os resultados empíricos apurados são articuladas com um conjunto abrangente de estudos anteriores, nacionais e internacionais, com eixos comuns a este, no sentido de enriquecer a discussão dos dados e identificar linhas de investigação futuras.

3.3. Hipóteses de trabalho

A exploração da pesquisa teórica e empírica nacional e internacional sobre a avaliação nas OTS, a partir de uma extensa revisão bibliográfica, permitiu sustentar a construção de um sistema de hipóteses orientadoras da pesquisa que visa dar conta dos padrões de avaliação adotados pelas organizações do terceiro setor no nosso país, ao longo das duas últimas décadas.

Hipótese 1 - Perante os novos desafios que se colocam às OTS decorrentes das mudanças nas exigências e discursos sobre desempenho e avaliação – a influência dos pressupostos da nova gestão pública, a crescente pressão para a profissionalização técnica

e da gestão, as novas exigências de qualidade e requisitos de maior eficácia e eficiência na intervenção social, etc. – é de esperar a prevalência de modelos de avaliação managerialistas, tecnocráticos e não participativos, orientados para a prestação de contas ascendente (a que é feita a agências financiadoras e reguladoras) em detrimento de padrões de avaliação participativos, pluralistas e centrados na aprendizagem organizacional.

- a. As entidades tendem a privilegiar sistemas de avaliação alicerçados em informação quantitativa sobre recursos e realizações face à avaliação qualitativa sobre as mudanças e impactos, concentrando-se sobretudo na conformidade com métricas, padrões e objetivos predeterminados.
- b. A adesão à avaliação nas OTS é primeiramente impulsionada por forças externas (coercivas, normativas e/ou miméticas), sobretudo as associadas ao financiamento, as quais exercem um forte controlo sobre o tipo de práticas de avaliação adotadas.
- c. Os fatores contingenciais internos, nomeadamente a dimensão, a antiguidade e as funções/papéis das OTS são fortemente influentes da capacidade avaliativa e do tipo de práticas e usos da avaliação.
- d. O enquadramento da avaliação em iniciativas de gestão mais amplas, como a gestão estratégica ou implementação de sistemas de qualidade contribui para conferir à avaliação uma estrutura mais formalizada e articulada com os processos de decisão de nível organizacional.
- e. Os processos de avaliação desligados da missão, que respondem a imperativos de curto prazo da prestação de contas externa, falham enquanto estímulo à introdução de mudanças positivas no comportamento organizacional.
- f. Uma cultura organizacional assente no compromisso da liderança com a avaliação e numa atitude positiva das pessoas envolvidas nos processos de recolha e análise da informação é decisiva na internalização da avaliação.

Hipótese 2 - Enquanto *stakeholders*-chave, as entidades financiadoras ou tutelares desempenham um papel fundamental na construção de processos de avaliação potenciadores da aprendizagem organizacional.

- a. Nas situações em que as entidades financiadoras privilegiam os padrões de prestação de contas, centrados nas metas quantitativas e execução física e financeira de curto prazo, que não capacitam as OTS, limitam a “responsabilidade sentida” e o uso da avaliação para a aprendizagem organizacional por parte das organizações que financiam.
- b. As situações em que as entidades financiadoras partilham com as OTS uma perspetiva de avaliação de longo prazo, e disponibilizam recursos financeiros, técnicos (formação, manuais, etc.) e de tempo necessários para a análise das lições aprendidas e para a reflexão sobre impactos do trabalho, potenciam o desenvolvimento de uma cultura de aprendizagem nas OTS beneficiárias.

Hipótese 3 - O FSE tem criado um ambiente, uma linguagem e um mercado de avaliação, contribuindo significativamente para o desenvolvimento de uma cultura de avaliação.

- a. O investimento do PIC EQUAL na promoção e organização de práticas de avaliação dos projetos (edição de manuais de apoio, realização de sessões de esclarecimento) destaca-se como o Programa do FSE que mais contribuiu para o desenvolvimento de uma cultura e avaliação nas organizações beneficiárias.
- b. O efeito potenciador do PIC EQUAL não tem sido sustentado pelos programas que lhe sucederam (caso do POPH).

3.4. As fontes e estratégias de recolha e análise da informação: um desenho metodológico misto

A prossecução dos objetivos da pesquisa implica a mobilização de um quadro metodológico capaz de articular, simultaneamente, uma abordagem extensiva facilitadora da caracterização das práticas da avaliação adotadas pelas OTS e respetivos contextos, com

uma abordagem intensiva que nos permita compreender os discursos e as interações protagonizadas pelos diversos atores implicados na construção de uma cultura de avaliação no terceiro setor.

Com efeito, a estratégia de pesquisa seguida baseou-se num desenho **metodológico misto** (Creswell, 1998), sustentado no paradigma pragmatista, cujos pressupostos foram explicitados no primeiro capítulo. A investigação orientou-se de acordo com os cinco propósitos essenciais dos estudos com configurações metodológicas mistas identificados por Greene, Caracelli e Graham (1989: 259): (i) a procura de convergência e corroboração através de diferentes métodos (*triangulação*); (ii) a procura de elaboração, ilustração e clarificação dos resultados de um dispositivo metodológico através dos resultados de outro (*complementaridade*); (iii) o uso dos resultados de uma metodologia para desenvolver ou informar outra (*desenvolvimento*); (iv) descobrir paradoxos e contradições que levam a um reenquadramento das questões de investigação (*iniciação*); (v) procurar expandir o alcance e a diversidade da investigação através do uso de diferentes métodos para diferentes componentes da pesquisa (*expansão*).

Combinaram-se exercícios de análise extensiva e intensiva nos diferentes eixos analíticos da investigação, triangulando informações provenientes das diversas fontes de informação. Ambos os tipos de análise favorecem uma confrontação dos dados recolhidos com diferentes técnicas, refletindo-se sobre os questionamentos sugeridos. **A abordagem intensiva** está presente na análise interpretativa das fontes documentais e dos registos orais das entrevistas semiestruturadas realizadas a 45 atores dos campos da avaliação e do terceiro setor. **O exercício metodológico de tipo extensivo**, por seu turno, consistiu na aplicação de um inquérito por questionário *online* a representantes de 575 organizações do terceiro setor. Este dispositivo permitiu não apenas obter informação extensiva sobre características, práticas e opiniões sobre avaliação nas OTS, quantificando uma multiplicidade de dados, mas também, sobretudo, relacionar entre si essa informação através de numerosas análises de correlação. As análises estatísticas bivariadas e multivariadas foram, nesta linha, decisivas para estabelecer padrões e tipologias de representações e práticas.

A força dos resultados das análises quantitativas foi associada com as descrições explanatórias aprofundadas obtidas através da análise qualitativa, sobretudo a conduzida a partir de entrevistas, mas também a partir dos diversos tipos de documentos analisados. A recolha de dados qualitativos foi estruturante do processo de investigação: além de ter sido crucial para delimitar e desenhar a fase quantitativa do estudo, carreando elementos essenciais para a seleção de questões e categorias de resposta do questionário às OTS, revelou-se também ser essencial na contextualização e interpretação dos resultados quantitativos. Por exemplo, foi na sequência de resultados das entrevistas com OTS, os quais foram reveladores de uma grande diversidade de enquadramentos dos exercícios avaliativos, que se optou por introduzir no questionário questões sobre um conjunto amplo de práticas que as organizações associam à medição/análise do desempenho, no sentido de captar essa diversidade e ponderá-la na análise. A combinação de métodos permitiu ainda expandir o alcance do estudo. Por exemplo, a pesquisa documental permitiu diversificar (e aprofundar) a análise de componentes da avaliação associadas ao ambiente externo, nomeadamente as relativas aos requisitos de monitorização e avaliação associados à tutela e financiamento das OTS e aos recursos financeiros e técnicos (formação) disponibilizados por entidades financiadoras e outros *stakeholders* às organizações.

A fertilidade desta triangulação metodológica traduz-se numa melhor compreensão dos processos sociais que estão subjacentes às práticas de avaliação das organizações que me proponho estudar, bem como na identificação dos mecanismos através dos quais estas se têm vindo a desenvolver.

Tendo em conta a escassez de investigações em Portugal sobre a temática da avaliação nas OTS, os dados produzidos foram analisados com base naquilo que eles próprios podem informar, renunciando a uma grelha concetual fechada e construindo e reconstruindo categorias à medida que as descobertas emergem, num caminho identificado com a abdução (Eco, 2016).

3.4.1. Entrevistas

A construção dos guiões orientadores das entrevistas foi realizada de acordo com o quadro teórico e analítico do trabalho, embora numa perspetiva aberta, de modo a dar

espaço à emergência de temas e categorias não previstas. Foram desenvolvidos quatro guiões de entrevistas semiestruturadas individuais (anexo 1) para cada um dos grupos de pessoas: a pessoal dirigente e técnico de OTS, representantes de organizações de cúpula do setor, especialistas e profissionais de avaliação e representantes de entidades financiadoras (dos Programas POPH e EQUAL). A natureza semiestruturada das entrevistas favoreceu a flexibilidade necessária ao ajustamento dos guiões à condição e experiência relevante em cada caso.

O trabalho de campo de realização das entrevistas foi desenvolvido entre Outubro de 2013 e Dezembro de 2014, período intervalado por pausas nas quais analisei o material entretanto produzido e persegui novas pistas de análise suscitadas pelos temas emergentes das narrativas dos sujeitos, carreando contributos para novas pesquisas bibliográficas e documentais e para a seleção de outras pessoas e organizações a entrevistar. Algumas entrevistas com profissionais de avaliação foram realizadas já depois daquele período, por dificuldades de agenda.

Os contextos de realização das entrevistas foram escolhidos tentando articular a conveniência das pessoas entrevistadas com a da pesquisa. Os diálogos foram conduzidos presencialmente (nos locais de trabalho das pessoas entrevistadas) ou por via telefónica⁵⁹

⁵⁹ As entrevistas por esta via acabaram por se revelar uma opção vantajosa para as entrevistas realizadas no norte e sul do país, não só por permitirem economia de tempo e recursos, mas também por tornarem mais ágil e flexível o agendamento das entrevistas, e por terem demonstrado ser profícuas em termos de capacidade heurística. Embora existam diferentes pontos de vista sobre a conveniência e utilidade de usar o telefone para realizar entrevistas, as suas vantagens são, em geral, reconhecidas, nomeadamente na exploração de temas mais sensíveis. A entrevista telefónica, além de minimizar as desvantagens associadas à entrevista pessoal (reduzindo enviesamentos nas respostas induzidos por expressões faciais ou pela deseabilidade social), associa-se a uma maior perceção de anonimato e tende a aumentar o nível de conforto para ambas as partes, potenciando uma entrevista mais relaxada, capaz de revelar informações sensíveis (Musselwhite *et al.*, 2006; Novick, 2008). Por outro lado, esta modalidade de entrevista apresenta menor vantagem, quer ao nível da clareza da comunicação, quer em termos de manutenção do envolvimento do/a entrevistado/a, ambos propiciados pelo contexto de interação pessoal, no qual podem ser usadas pistas visuais (Garbett e McCormack, 2001). Neste estudo, estas particularidades de interação não parecem ter suscitado diferenças significativas nos resultados obtidos. As respostas abrangentes suscitadas pelas pessoas entrevistadas por via telefónica sugerem que esta modalidade pode gerar informação detalhada e adequada ao tipo de pesquisa qualitativa. Análises anteriores haviam já evidenciado poucas diferenças nos tipos de resposta obtidas em entrevistas telefónicas e face-a-face. Por exemplo, Sturges e Hanrahan (2004: 113) concluíram que não havia diferenças visíveis entre as respostas dadas nas duas modalidades, e que havia "similaridades na quantidade, natureza e profundidade das respostas". No mesmo sentido apontam os estudos de Sempik *et al.* (2007) e Bryman *et al.* (2008), levando um dos autores a concluir que quando se considera a economia de tempo e de custos de deslocação, uma vez que é possível abranger respondentes

e assentaram em questões incluídas em guiões previamente definidos⁶⁰. Todas as pessoas entrevistadas aceitaram a gravação áudio das entrevistas o que facilitou a fixação do seu conteúdo para posterior transcrição e análise. A receptividade à realização das entrevistas foi genericamente boa, existindo empatia relativamente ao objeto de estudo, pela adesão à ideia da relevância da avaliação. As maiores dificuldades foram sentidas na mobilização de representantes de organizações de cúpula (dificuldades de agenda ou não resposta aos convites para entrevista) que, em alguns casos, não foram superadas. A duração das entrevistas variou de acordo com a disponibilidade das pessoas, sendo que a duração mais comum oscilou entre uma hora e uma hora e meia⁶¹.

A análise de conteúdo das entrevistas assentou no procedimento de codificação temática com recurso a categorias pré-estabelecidas de acordo com o quadro teórico e outras que emergiram indutivamente dos discursos dos atores. Uma análise rigorosa, baseada em sucessivas leituras do material, permitiu-me criar, em diálogo com os conceitos e hipóteses iniciais, categorias e estruturas abstratas. Esta grelha analítica permitiu-me organizar os conteúdos das entrevistas. Para a análise horizontal recorri a um programa informático de análise qualitativa – o programa MAXQDA (Qualitative Data Analysis Software), uma aplicação de análise qualitativa de dados e métodos mistos que permite uma fácil integração com outras ferramentas de análise de dados. O texto da transcrição das entrevistas em formato *Word* foi transferido para uma base documental no MAXQDA a partir da qual foram indexados extratos de cada uma das entrevistas ao sistema de categorias (vide anexo 3) produzido em diálogo com o complexo teórico de partida e as ideias e relações estabelecidas pela análise da informação produzida. Apesar da intensidade e morosidade do processo de codificação das entrevistas, este programa permitiu um trabalho de organização da informação precioso no esforço de análise e reflexão. As suas vantagens em termos de manuseamento, organização, articulação e

amplamente dispersos pelo território, o método poderá ser considerado altamente eficiente, face ao grande volume de informação recolhida (Bryman, 2012).

⁶⁰ Os guiões de entrevista foram elaborados como roteiro de condução das conversações e, além de um núcleo central de questões orientadoras comuns a todas as pessoas entrevistadas, foram adaptados à especificidade de cada tipo de organização/interlocutor/a.

⁶¹ A entrevista mais curta teve a duração de trinta minutos e a mais longa de duas horas e quarenta e cinco minutos.

leitura categorial da informação são claras, permitindo articular uma leitura vertical e horizontal dos dados (Monteiro, 2011).

Entrevistas focalizadas individuais com dirigentes e/ou técnicos de OTS

Foram conduzidas entrevistas com 40 representantes⁶² de 31 organizações de natureza, dimensão (financeira e de recursos humanos) e localização diversa (cf. listagem de entrevistas no anexo 2). Tendo em conta a composição heterogénea do terceiro setor português, a seleção da amostra de organizações a entrevistar teve por base a decisão de incluir vários tipos de entidades (cooperativas, IPSS, ADL, ONGD, fundações, associações empresariais e sindicais), de diferentes dimensões em termos financeiros e humanos (pequenas e grandes organizações), e de diferentes regiões do país. Teve também a preocupação de incluir organizações com projetos POPH e/ou EQUAL, considerando a relevância analítica atribuída a estes programas. A tabela que se segue sintetiza as principais características das organizações entrevistadas.

Tabela 1. Variáveis de caracterização das OTS entrevistadas

		N	%
Tipo de organização ⁶³	ADL	5	16,1
	IPSS	12	38,7
	ONG	11	35,5
	Outra*	3	9,7
Funções ⁶⁴	Expressão	3	9,7
	Serviços	20	64,5
	Serviços+ Expressão	8	25,8

⁶² Em entrevistas com 8 organizações participou mais do que uma pessoa (normalmente 2). No convite endereçado às organizações foi solicitado o “agendamento de entrevistas individuais com dirigente ou técnico/a que tivesse vindo a acompanhar ou fosse conhecedor/a das práticas da organização ao nível da gestão e avaliação e com pessoal técnico (uma ou duas pessoas) com envolvimento em projetos POPH/EQUAL.” Em geral, os/as responsáveis das organizações entenderam disponibilizar-se ou indicar apenas uma pessoa para ser entrevistada. Nos casos em que consideraram adequado mobilizar mais do que uma pessoa para ser entrevistada assumiu-se que a participação na entrevista seria conjunta.

⁶³ As organizações foram agrupadas consoante a sua forma jurídica, estatuto e vocação principal (vide secção 2.2).

⁶⁴ Tipologia adotada do projeto comparativo do Center for Civil Society Studies da Universidade de Johns Hopkins, construída a partir das principais áreas de atividade das organizações (vide secção 2.2).

Antiguidade	Até 20 anos	12	38,7
	21 a 50 anos	13	41,9
	Mais 50 anos	6	19,4
Dimensão financeira	Pequena (Até 99.999€)	6	19,4
	Média (100.000€ a 1.000.000€)	11	35,5
	Grande (+1.000.000€)	14	45,2
Pessoal Renumerado	1 a 5 pessoas	6	19,4
	6 a 25 pessoas	8	25,8
	26 a 50 pessoas	4	12,9
	51 a 100 pessoas	7	22,6
	Mais de 100 pessoas	6	19,4
Região	Sul	4	12,9
	Centro	15	48,4
	Lisboa e Vale Tejo	5	16,1
	Norte	7	22,6

Fonte: Inquérito por questionário.

Nota: * Uma cooperativa de habitação, uma cooperativa de ensino e uma escola de formação.

A maioria das pessoas participantes nas entrevistas ocupava posições de direção ou coordenação na organização: 18 pertenciam à direção estatutária, 15 à direção/coordenação técnica/executiva, e 6 eram outros/as técnicos/as.

Entrevistas focalizadas individuais com dirigentes de plataformas representativas de OTS

Foram contactados para a realização de entrevista dirigentes das maiores organizações de cúpula do terceiro setor (CNIS, União das Mutualidades, União das Misericórdias, ANIMAR, FENACERCI, CPCCRD, CONFECOOP, Plataforma das ONGD). Apesar dos reiterados convites, acederam ao meu convite para entrevista representantes de quatro estruturas: CNIS, FENACERCI, CPCCRD e Plataforma das ONGD.

Entrevistas focalizadas individuais com especialistas e profissionais de avaliação

Foram realizadas 7 entrevistas com especialistas e consultores/profissionais de avaliação em âmbitos diversos: avaliação de políticas e programas públicos, avaliação de projetos de intervenção social promovidos por OTS e avaliação organizacional.

Entrevistas focalizadas individuais com responsáveis pelos órgãos de gestão e de acompanhamento do PIC EQUAL e do POPH (entidades financiadoras)

Foram realizadas entrevistas com elemento da Direção da Comissão Executiva do POPH, da Direção do Observatório do QREN, e com responsável do Gabinete de Gestão do PIC EQUAL.

3.4.2. Inquérito por questionário às organizações do terceiro setor

O inquérito, disponível em anexo (vide anexo 4), foi estruturado em várias áreas, correspondentes aos objetivos do estudo, e elaborado em diálogo com o tratamento de informação fornecida quer pelo *corpus* documental, quer pelas entrevistas, de modo a que fosse possível configurar de forma adequada as questões a inserir e as hipóteses de respostas.

Sendo objetivo do estudo abranger a globalidade das OTS (e não apenas entidades com projetos EQUAL e/ou POPH), na ausência de uma base de dados consolidada que correspondesse ao universo das OTS e na qual constassem os respetivos contactos, recorri a uma multiplicidade de fontes que me permitisse dar conta da diversidade que atravessa o terceiro setor em Portugal. A base de contactos do universo das OTS – que totalizou 2.897 organizações com endereços de *email* válidos – foi construída através de: listagens de organizações com endereço de correio eletrónico disponível em páginas *web* (associadas da união das misericórdias, associadas da união das mutualidades, ONGD registadas junto do Instituto Camões, ONG inscritas na Fundação Calouste Gulbenkian); listagens disponibilizadas por estruturas representativas do setor (Cooperativas, pela CASES, CERCIS, pela FENACERCI, ADL, pela ANIMAR, Coletividades, pela CPCCRD). Uma vez que a CNIS não disponibilizou os contactos das IPSS suas associadas, foi-lhe solicitado que encaminhasse diretamente para elas o meu *email* com o convite às organizações para responder ao

inquérito. Todavia, a CNIS recusou encaminhar o *email*, incluindo o meu convite às organizações na sua newsletter semanal, o que limitou enormemente o efeito pretendido.

Desenvolveu-se um processo de inquirição em plataforma de internet através do recurso à aplicação *LimeSurvey*. Todas as 2.897 OTS da base de sondagem foram convidadas, por *email*, a participar no preenchimento do questionário. O questionário foi aplicado entre 26 de Maio e 20 de Junho de 2014, tendo antes sido realizados vários pré-testes presenciais.

Depois de um processo de algumas insistências que se prolongou por cerca de 3 semanas, foram recebidas respostas de **575 organizações**, tendo sido atingida uma **taxa de resposta de cerca de 20%**, o que é significativo para esta modalidade de aplicação de questionários. O processo resultou numa amostra heterogénea de organizações, integrando os diferentes tipos de Organizações do Terceiro Setor, seja na diversidade de formas e estatutos jurídicos, seja na multiplicidade de âmbitos de atividade, funções, dimensões, antiguidade e distribuição geográfica.

Para o tratamento das respostas obtidas foi utilizado o SPSS – *Statistical Package for the Social Sciences* – tendo recorrido a análises estatísticas descritivas e inferenciais, univariadas, bivariadas e multivariadas: análises de frequência, análises de contingência, associação e correlação, testes paramétricos e não paramétricos, modelos de regressão e análises exploratórias multivariadas (Marôco, 2010; Pestana e Gageiro, 2016).

Um conjunto de questões de caracterização das organizações permite desde já definir as características da amostra e, a partir de algumas delas, estimar a sua aproximação ao universo do terceiro setor em Portugal. Foi tendo em conta a composição do terceiro setor português, que o estudo quantitativo nesta investigação contemplou uma amostra heterogénea de organizações, integrando as diferentes famílias do terceiro setor, seja na diversidade de formas e estatutos jurídicos, seja na amplitude de funções, idades, localizações geográficas, número de colaboradores/as e orçamentos.

Embora não se possa falar de representatividade da amostra, dado que se trata de uma amostra não probabilística, composta por voluntários/as, a análise da tabela 2 mostra que se encontra abrangida no estudo a diversidade de organizações do setor, quer no que toca à forma e estatuto jurídico, quer no que se refere às áreas de atividade.

Tabela 2. Tipologia das organizações da amostra, por comparação com a Conta Satélite da Economia Social (CSES)⁶⁵ (%)

		Cooperativas	Mutualidades	Misericórdias	Fundações	Associações e outras OES	Total
CSES	OTS	3,5	0,2	0,6	0,9	94,8	61.268
	IPSS	2,4	1,9	6,9	3,9	84,7	5.584
Amostra	OTS	8,2	1	12,2	3,3	75,3	575
	IPSS	11,3	2,4	19	5,7	61,5	247

Fonte: Inquérito por questionário.

Quando comparada com os dados da Conta Satélite da Economia Social (CSES), a estrutura da amostra no que se refere à tipologia das organizações não diverge muito, revelando um peso elevado da categoria associações e outras organizações do TS (tabela 2). Todavia, a proeminência desta categoria é menos marcada neste estudo, devido ao forte enviesamento da amostra na área da ação social, que se distingue pela maior presença de cooperativas de solidariedade social e misericórdias. Por outro lado, apesar de apresentarem na amostra um peso inferior ao evidenciado na estrutura global do setor, o associativismo de base popular (coletividades de cultura, recreio e desporto) encontra-se vigorosamente representado, constituindo o segundo grupo de entidades mais retratado no estudo, abrangendo 27% da amostra, como se vê na tabela abaixo.

Tabela 3. Variáveis de caracterização das OTS inquiridas

		N	%
Tipo de organização	ADL	37	6,4%
	IPSS	234	40,7%
	ONG	105	18,3%
	Coletividades	155	27,0%
	Outras	44	7,7%
Funções	Serviços	260	45,2%
	Expressão	170	29,6%
	Serviços + Expressão	145	25,2%
Antiguidade	Até 20 anos	189	34,6%
	21 a 50 anos	213	38,9%
	51 ou mais	145	26,5%

⁶⁵ INE e CASES (2016).

Dimensão financeira	Pequenas (até 99.999€)	272	51,4%
	Médias (100.000€ a 1.000.000€)	143	27,0%
	Grandes (+ 1.000.000€)	114	21,6%
Pessoal remunerado	Nenhuma	93	16,2%
	1a 5	111	19,3%
	6 a 10	62	10,8%
	11 a 25	97	16,9%
	26 a 50	86	15,0%
	51 a 100	72	12,5%
	Mais de 100	50	8,7%
	NR	4	0,7%
Região	Norte	147	25,6%
	Centro	132	23,0%
	Lisboa e Vale do Tejo	214	37,2%
	Sul	76	13,2%
	Ilhas	5	1%
	NR	1	0,2%

Fonte: Inquérito por questionário.

Responderam ao inquérito OTS de todas as regiões portuguesas (NUTS II), à exceção da região autónoma da Madeira. A maioria das entidades localiza-se na zona centro do país, sobretudo na região de Lisboa e Vale do Tejo, a que se segue a região Norte, seguida a alguma distância da região Sul.

Coerentemente com o perfil da atividade das organizações inquiridas, as entidades prestadoras de serviços são prevaletentes na amostra, abrangendo mais de dois terços das entidades, um quarto das quais acumulam funções de expressão. São menos de um terço as organizações com funções eminentemente expressivas, grupo que é composto sobretudo por coletividades. A par com as IPSS, este tipo de entidade destaca-se também pela maior antiguidade, contrapondo com as ADL e ONG, de constituição mais recente.

Em termos de dimensão, aqui aferida a partir do volume do orçamento, destaque-se a forte presença de entidades de muito pequena e pequena dimensão financeira, que abrangem cerca de metade da amostra, e são compostas sobretudo por coletividades (seguidas das ONG). No antípoda encontramos as IPSS: mais de dois terços têm uma grande dimensão financeira. As ADL, por seu turno, destacam-se entre as organizações de dimensão intermédia. A estrutura em termos de dimensão financeira é semelhante à do pessoal remunerado, com perto de metade das entidades a contar com 10 ou menos

trabalhadores/as remunerados/as (sobretudo coletividades e ONG), e pouco mais de 20% (sobretudo IPSS e *outras* entidades) a integrar mais de meia centena de pessoas remuneradas.

3.4.3. Análise documental

A análise documental foi um importante elemento do trabalho de pesquisa que visou sobretudo consolidar informação quer sobre requisitos, ferramentas e expectativas de avaliação associadas quer ao ambiente externo (requisitos e recursos de avaliação associados aos financiamentos – em particular do POPH e PIC EQUAL), e iniciativas formativas em avaliação, quer às próprias práticas de avaliação das organizações.

Foi criada uma base de dados documental sobre fundos estruturais europeus – e outras fontes de financiamento relevantes – e avaliação, com o material significativo disponível: legislação, regulamentação e documentos de política; relatórios de execução e avaliação, guias, manuais e orientações dos programas; guias metodológicos e outros materiais de apoio; formulários e *templates*. A base documental foi construída, por um lado, a partir do acesso às *páginas web* dos quadros comunitários, programas, entidades e estruturas de coordenação, gestão e monitorização do FSE, e, por outro lado, às páginas de outras entidades financiadoras (Segurança Social, Institutos Camões, ACM, Fundação Calouste Gulbenkian, Fundação EDP)/linhas de financiamento (Programa Escolhas, CLDS, Cidadania Ativa, Cooperação/Educação para o Desenvolvimento) destacadas nas entrevistas como relevantes no domínio da avaliação. Esta base integrou também os referenciais de gestão/avaliação da qualidade (manuais, guias de certificação, modelos de avaliação, etc.) associados à certificação da formação (Direção Geral do Emprego e das Relações de Trabalho) e da qualidade (ISO 9001, Modelo de Avaliação da Qualidade das Respostas Sociais – ISS e Modelo da Qualidade Europeu para os Serviços Sociais – EQUASS), aludidos amiúde enquanto enquadradores dos processos de avaliação.

Uma segunda base documental, que foi sendo alimentada pelos contributos das entrevistas realizadas com os diversos *stakeholders*, integrou instrumentos de capacitação em avaliação para OTS: ações formativas, com destaque para as conduzidas no âmbito do Programa de Formação-Ação para Entidades da Economia Social (medida 3.1.2 do POPH), e outros recursos técnicos no domínio da avaliação (plataformas de partilha de

metodologias e instrumentos, guias, manuais, ferramentas, etc.). A base documental de iniciativas formativas em avaliação foi construída a partir de pesquisa exploratória conduzida *online* no âmbito deste estudo. Além da informação recolhida a partir do inquérito às OTS e das entrevistas com OTS, organizações de cúpula, especialistas/profissionais de avaliação e representantes de entidades financiadoras, a estratégia para fazer o levantamento da oferta formativa em avaliação nas OTS consistiu na pesquisa nas páginas web de entidades ou plataformas identificadas nas entrevistas como relevantes em termos de promoção e/ou divulgação de ações de capacitação em avaliação para o terceiro setor, com destaque para entidades representativas do setor, mas também outras redes e plataformas: CNIS – Confederação Nacional das Instituições de Solidariedade; União das Misericórdias Portuguesas (UMP); União das Mutualidades (UMP); Associação Portuguesa para Desenvolvimento Local (ANIMAR); Federação Minha Terra – Federação Portuguesa de Associações de Desenvolvimento Local; Centro Português de Fundações (CPF); Confederação Portuguesa das Coletividades de Cultura, Recreio e Desporto (CPCCRD); Federação Nacional de Cooperativas de Solidariedade Social (FENACERCI) e Plataforma Portuguesa das ONGD. A procura nestes sites foi feita, sempre que possível, a partir dos motores de busca das próprias páginas, mediante a introdução, à vez, dos vocábulos “avaliação”, “desempenho” e “impacto”. Nos casos em que estes motores não estavam disponíveis nas páginas, a pesquisa foi feita a partir dos respetivos menus/submenus. Paralelamente, foi conduzida, no dia 21 de Abril de 2016, uma pesquisa na Internet, através do motor de busca da Google (Google.pt), sendo considerados os resultados desde Janeiro de 2010. Foi utilizado o modo “Pesquisa Avançada” para introduzir os seguintes parâmetros de busca: No campo “todas estas palavras”, foram inseridas, à vez, as palavras “avaliação” e “impacto”; no campo “exatamente esta palavra ou expressão” foram incluídas, também à vez, as expressões “terceiro setor”, “economia social”, “economia solidária”, “sem fins lucrativos”, “empreendedorismo social”, “projetos sociais”, no campo “qualquer uma destas palavras”, foram introduzidas as palavras “formação”, “workshop”, “curso”, “unidade curricular”, “disciplina”; “mestrado”, “doutoramento”, “licenciatura”, “graduação”, “especialização”, “seminário”, “congresso”, “conferência” e “palestra”; no campo “região” foi introduzida a palavra “Portugal”; no

campo “termos que aparecem”, foi escolhida a opção “no texto da página”. Ao todo foram geradas 12 pesquisas diferentes, já que nos campos “todas estas palavras” e “exatamente esta palavra ou expressão” os termos foram introduzidos à vez. Foram contabilizadas as diversas edições de um mesmo evento que se repete, de forma a captar a intensidade da oferta formativa.

A terceira e última base documental foi construída a partir da disponibilização pelas organizações entrevistadas de documentos que sistematizassem exercícios avaliativos conduzidos pelas entidades nos últimos 3 anos: relatórios/planos de atividades⁶⁶, planos estratégicos, relatórios de avaliação de projetos/de iniciativas/organizacionais, etc. Das 31 organizações auscultadas via entrevista, responderam à minha solicitação – ou tinham disponíveis alguns destes documentos nos respectivos sítios da internet – 21 entidades⁶⁷, que remeteram 48 documentos, no total. A maioria das entidades optou por enviar apenas os relatórios de atividades anuais (frequentemente acompanhados de planos anuais), sob o fundamento de estes sintetizarem a generalidade dos exercícios de avaliação conduzidos pela entidade ao longo do ano. Entre os outros documentos enviados incluem-se: relatórios de avaliação de projetos (4 entidades), planos estratégicos (3 entidades), relatórios de avaliação da satisfação de *stakeholders* (3 entidades), relatório de avaliação de impacto organizacional (1 entidade), relatório de revisão/acompanhamento de entidade certificadora (1 entidade), relatório com dados das respostas sociais (1 entidade), revista da organização (1 entidade), dissertação de licenciatura de estagiário (1 entidade).

Dada a amplitude, diversidade e heterogeneidade encontrada dentro de cada uma das bases documentais, a informação documental foi sujeita a uma análise intensiva não estruturada e essencialmente vertical. Foi possível, não obstante, fazer leituras horizontais em algumas dimensões de análise da informação produzida a partir da segunda e terceira bases documentais. Por exemplo, na base documental de iniciativas formativas em avaliação foi possível fazer leituras horizontais sobre a duração, conteúdo e tipo de

⁶⁶ No caso dos relatórios e planos de atividades anuais foram requeridos apenas os do ano transato.

⁶⁷ Apesar das insistências (por email e telefone) 10 das entidades entrevistadas não acederam a remeter qualquer documento, nem este tipo de documento (inclusive relatórios e planos de atividades) constava nos respetivos *sites* institucionais.

entidade promotora. Na base de documentos disponibilizados pelas OTS foi também possível fazer algumas análises transversais sobre o tipo de informação de avaliação produzida (qualitativa/quantitativa, descritiva/crítica, sobre inputs/outputs/efeitos).

PARTE II: O ESTUDO EMPÍRICO

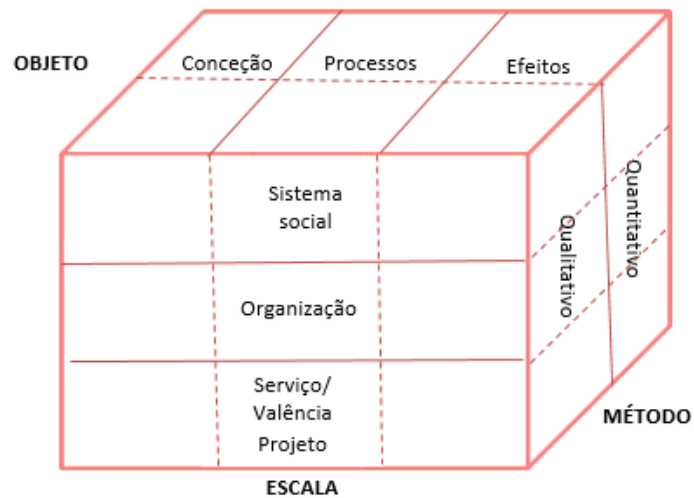
Capítulo 4. Discursos e práticas de avaliação no terceiro setor português

Neste capítulo serão exploradas as modalidades de monitorização e avaliação adotadas pelo terceiro setor em Portugal, partindo do olhar sobre seu desenho metodológico e sobre as lógicas discursivas inerentes às estratégias prosseguidas. A análise incidirá também nos fatores explanatórios da adoção de práticas avaliativas, nos seus propósitos e na forma como é usado o conhecimento produzido pela avaliação. Este exercício é conduzido a partir da informação recolhida através do inquérito às organizações do terceiro setor, de entrevistas conduzidas quer com responsáveis destas entidades e com representantes de organizações de cúpula do setor, quer com especialistas/profissionais de avaliação e dos documentos disponibilizados pelas organizações.

Refletidas numa grande diversidade de análises (por exemplo, Osborne, 1992; Tassie, Murray e Cutt, 1998; Capucha *et al.*, 1996; Monteiro, 1996; Cutt e Murray, 2000; Stufflebeam, 2014; Stufflebeam e Coryn, 2014), as modalidades que os processos de avaliação podem assumir são muito variáveis quanto aos objetivos, intervenientes, à dimensão dos objetos sobre os quais incidem e respetivas características, aos públicos-alvo, às estratégias e objetivos e aos produtos e metodologias. Cada avaliação tem um âmbito distinto em função das perguntas a que procura responder, o seu propósito, natureza, método, fase em que se aplica, consideração dos/as utilizadores/as, ou utilidade.

A figura 4 ilustra as principais dimensões em que se estrutura a análise das modalidades dos processos de avaliação nas organizações estudadas, e pretendem responder essencialmente às questões “o que é avaliado” e “como é avaliado”.

Figura 4. Enquadramento do processo de avaliação



Fonte: Adaptado de Cutt, James; Murray, Vic (2000) *Accountability and Effectiveness Evaluation in Non-Profit Organizations*. Abingdon: Taylor & Francis, Figura 3.1.

A primeira dimensão do enquadramento adotado – **escala de avaliação** – remete para a “entidade” que está a ser avaliada. A avaliação pode envolver, pelo menos, 3 níveis de análise: micro (um projeto ou serviço/unidade organizacional)⁶⁸, meso (a organização como um todo) e macro (o sistema social mais alargado). O primeiro nível centra-se em elementos específicos no seio de uma organização, como seja um projeto, uma unidade organizacional (departamento, serviço, valência) ou uma função (recursos humanos, marketing, etc.). As OTS são compostas por diversas componentes concebidas para atingir a missão global da organização. Algumas avaliações centram-se no desempenho global da organização, na prossecução da sua missão, e menos no desempenho de cada um dos seus projetos ou respostas sociais. Finalmente, qualquer necessidade ou problema social é geralmente abordado por várias organizações e programas. O último nível de avaliação – avaliação do sistema social - analisa como os esforços de várias partes interessadas (às vezes chamados de "atores") têm um impacto coletivo sobre determinada questão.

A segunda dimensão – **objeto da avaliação** – remete para o âmbito de incidência ou conteúdo da análise avaliativa. Essencialmente, a avaliação nas organizações pode centrar-se na conceção (contexto e *inputs*), nos processos (atividades e realizações) e nos

⁶⁸ A avaliação pode ainda ocorrer a nível individual (mediante avaliação de desempenho individual). Todavia, a análise deste enfoque não é central neste estudo.

efeitos (mudanças). O primeiro âmbito de incidência remete, em primeiro lugar, para a análise dos contextos, mediante identificação de necessidades e problemas e, em segundo lugar, para o processo de formulação da intervenção, proporcionando informação sobre os recursos necessários para pôr em curso os processos. A avaliação de processos centra-se nos aspetos relativos à operacionalização e implementação das intervenções, incluindo a forma como os serviços são prestados e as práticas de gestão. O foco é sobretudo na compreensão da forma como as intervenções produzem os resultados e menos nos resultados em si. Nas avaliações com este foco cabe também a análise das realizações, isto é, dos níveis de execução das atividades previstas e os resultados de tipo mais imediato – e tangível (i.e., produto material, ou *output*, gerado diretamente pelas atividades). Por fim, a avaliação dos efeitos remete para as mudanças – eminentemente intangíveis – nos *stakeholders* ou comunidade global, esperadas ou não, direta ou indiretamente, suscitadas pela intervenção das organizações. Partindo da distinção entre a *qualidade do serviço* e a *qualidade de vida*, Osborne (1992) nota que o centramento neste objeto é o único que permite compreender os efeitos da intervenção na qualidade de vida das pessoas, enquanto a qualidade dos serviços humanos/sociais é passível de ser analisada a partir do foco nos recursos e nos serviços que as organizações prestam (i.e., nos *inputs* e nos *outputs*).

A última dimensão de análise – **método de avaliação** – respeita à orientação metodológica dos exercícios de avaliação. Para a análise desta dimensão é trazida a distinção fundamental entre abordagens qualitativas e quantitativas, nas suas diferentes tradições em termos de produção e de análise de informação, e de questões filosóficas adjacentes.

Uma observação prévia à análise dos processos de avaliação nas diversas dimensões de análise, e que decorre da análise da informação produzida no âmbito do estudo, respeita ao **conceito de avaliação**. Como bem notam King e Stevahn (2013:44), “o termo ‘avaliação’ significa diferentes coisas para diferentes pessoas” e por isso pessoas e organizações fazem avaliação de diferentes formas e com diferentes propósitos. Quando em contexto de entrevista se perguntou como é que as OTS conduziam as suas avaliações,

foi frequente a identificação de um conjunto de processos informais deajuizamento sobre o desempenho por parte das pessoas entrevistadas. À semelhança do que se observou noutras análises (Hall *et al.*, 2003; Murray, 2005; Carman, 2007), tornou-se evidente que a generalidade das organizações entrevistadas conduz uma grande diversidade de atividades, desde práticas informais e reativas, até exercícios mais sistemáticos e pró-ativos. Além disso, nas organizações tende-se a identificar como avaliação um conjunto de atividades de gestão que se associam ao que autores como Carman (2009) e Murray (2005) designam de “movimento da *accountability*”. Essas atividades de gestão incluem o planeamento estratégico, sistemas de gestão da qualidade, auditorias, avaliação individual do desempenho e o reporte (relatórios de contas ou de projetos, por exemplo).

Existe, pois, entre os/as participantes no estudo, uma perspetiva pouco rigorosa sobre o conceito de avaliação, sendo frequente a consideração de que as abordagens informais e contínuas de monitorização e apreciação do desempenho são tão ou mais importantes do que as abordagens mais formais e científicas. O excerto abaixo, extraído de entrevista com técnico de uma OTS é ilustrativo deste ponto de vista:

Quase sempre e só reflexão, quer de grupo quer individual. (...) Conversamos normalmente de 15 em 15 dias, mas nunca fazemos a ata das coisas. Há uma avaliação efetiva, não é assim tão informal como isso, mas é muito curta e com objetivos muito concretos: está a correr bem, as pessoas estão a ir, o formador é certo ou não para aquele grupo (...). E alteramos as nossas posturas e as nossas decisões dependendo daquilo que formos conversando. Quando não tivemos estas ações de formação, mas tivemos outras atividades com jovens, a avaliação era também informal mas era quase diária. (Entrevista com representante de ONG – 31)

Para propósitos da pesquisa foi, por isso, importante definir a avaliação de forma flexível e diversa, de molde a apreender a multiplicidade de práticas conduzidas pelo setor, abdicando (parcialmente) das definições encontradas no campo formal da pesquisa em avaliação, as quais assentam essencialmente nas metodologias de investigação social (Rossi, 2013). Assim, a definição de avaliação adotada inclui atividades como avaliação de programas/projetos, análises de satisfação (de clientes/utentes/participantes, membros e outros *stakeholders*), análises de impacto, diagnóstico organizacional, medição dos desempenhos (indicadores de desempenho) e outras atividades que apoiem as organizações do terceiro setor a ajuizar sobre o seu desempenho.

4.1. Tipos de discurso sobre avaliação: científico, associativo e gestor

As OTS oferecem um enquadramento excepcional para o cruzamento de diferentes lógicas e linguagens sobre avaliação produzidas nos diversos campos de estudo, já que o terceiro setor tem sido descrito como híbrido, misturando diferentes princípios e racionalidades, em virtude de ocupar uma posição intermédia dentro de um espaço delimitado por um triângulo cujos vértices são o mercado, o Estado e a comunidade (Evers e Laville, 2004). Ao interagirem com as instituições e racionalidades nos três vértices, são influenciadas por umas e outras, podendo assumir características variadas e com maior ou menor proximidade às características das organizações dos restantes “setores” (*ibidem*).

A discussão sobre avaliação que se faz no terceiro setor é, naturalmente, permeável às diversas racionalidades geradas no mercado, no Estado e na comunidade. No caso particular da avaliação e do desempenho organizacional, as lógicas do mercado e do Estado têm vindo a convergir, na sequência de reformas neoliberais e da orientação para a Nova Gestão Pública nos Estados ocidentais (por exemplo, Osborne e Gaebler, 1993), assentes na filosofia da alegada forte capacidade que a gestão empresarial tem para dar respostas aos problemas do mundo atual. Têm raízes na cultura da gestão de mercado práticas de avaliação associadas, designadamente, à gestão e planeamento estratégicos, aos sistemas de certificação da qualidade, ao controlo e auditoria e à avaliação de desempenho. A cultura associativa prevalecte na organização coletiva da sociedade civil ou das comunidades, por seu turno, associa à avaliação ideias como a participação, a mudança social, o “empowerment”, a missão social, os valores e a justiça. Além dos domínios sociais/culturais da sociedade civil e da gestão (empresarial e governamental), também o mundo da ciência tem desempenhado um papel relevante na história do desenvolvimento do terceiro setor e influencia os discursos e as práticas de avaliação nestas organizações. O domínio da ciência inclui práticas e discursos característicos da avaliação de programas e projetos, que tradicionalmente preconiza a utilização dos métodos da investigação em ciências sociais. **As lógicas de avaliação das organizações do terceiro setor situar-se-ão, nesta linha, num campo híbrido entre os discursos científico, da gestão e da sociedade civil** (Powell *et al.*, 2012, 2014, 2015).

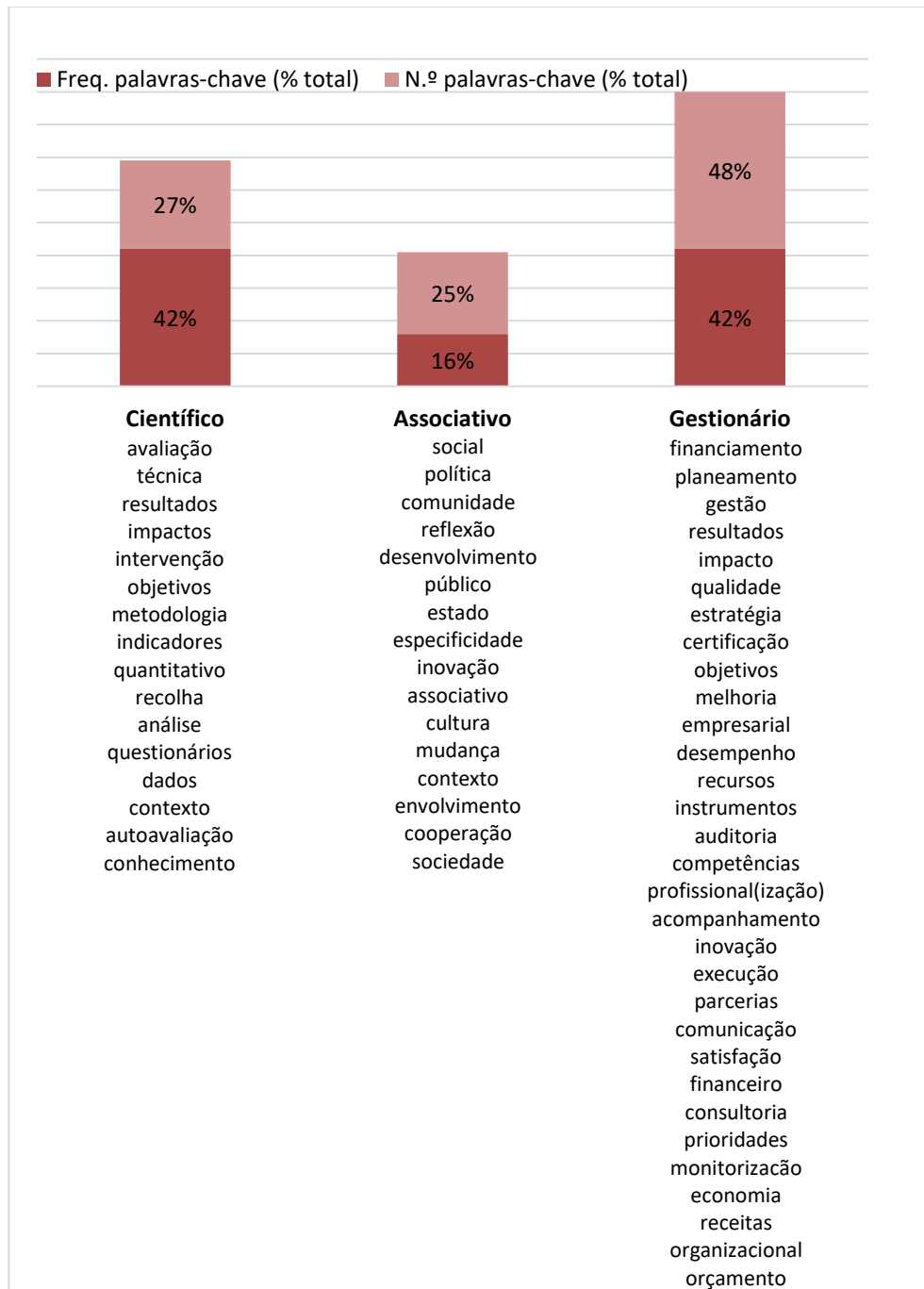
Além da informação substantiva sobre as lógicas inerentes às práticas e às percepções sobre avaliação nas OTS possibilitada pela análise de conteúdo das entrevistas realizadas, a análise do sentido do discurso das pessoas entrevistadas é também elucidativa do tipo de orientação semântica das narrativas sobre avaliação. Sendo a linguagem um elemento constitutivo da vida social, criadora de coesão e impositora de limites sociais, ela reflete e reforça sistemas normativos comuns, servindo para integrar e excluir e para criar comunidades (e.g., Brown e Yule, 1983). As palavras-chave, especificamente, constituem "termos significativos, indicativos de certas formas de pensamento" que compõe um vocabulário específico de um domínio distintivo (Williams, 1969 *apud* Powell *et al.*, 2012).

Assim, a partir da contagem das ocorrências de palavras no conjunto das entrevistas, foi selecionado um conjunto de termos e expressões comumente utilizados na literatura nacional e internacional sobre desempenho e avaliação nas organizações do terceiro setor. Apesar de subordinada ao *background* teórico alimentado pela pesquisa bibliográfica, o procedimento de seleção dos termos e expressões foi aberto, tendo-se optado por criar um conjunto de termos induzido a partir dos dados textuais. Dentre os termos relevantes na discussão sobre avaliação referenciados, excluíram-se os que são mencionados menos de 50 vezes no conjunto das entrevistas. Esta abordagem produziu uma lista de 57 palavras-chave, que foram categorizadas como tendo raízes na cultura científica, de gestão ou associativa. Por estar consciente do desafio que representa esta divisão, dada a fluidez das fronteiras dos três tipos de discurso, ressalvo que não se trata de uma divisão rígida ou estanque. No caso de alguns termos, a dificuldade em situá-los em apenas um dos domínios culturais levou-me a integrá-los em mais do que um. No domínio científico foram incluídos termos⁶⁹ com raízes na cultura científica, tais como "avaliação", "indicadores", "dados" ou "métodos". Já palavras como "qualidade", "certificação", "auditoria" e "satisfação" foram incluídas no domínio da gestão. Por fim, termos com raízes na cultura associativa prevalecte no terceiro setor, tais como "social", "mudança", "comunidade" e "desenvolvimento", foram incluídos no domínio associativo.

⁶⁹ Na generalidade dos casos, optou-se por agregar palavras da mesma família de palavras, isto é, com origem no mesmo radical ou palavra, e portanto relacionadas morfológica e etimologicamente. Por exemplo, o termo "avaliação" agrega outros da mesma família de palavras: avaliar, avaliador, avaliado.

A análise do gráfico abaixo mostra-nos que embora os termos que refletem valores associativos tenham relevância no discurso sobre avaliação, a **linguagem que indica uma orientação científica e de gestão é prevalecente**.

Gráfico 1. Tipos de discurso sobre avaliação: científico, associativo e gestor
(N= n.º de ocorrências)



Fonte: Entrevistas semiestruturadas.

Nota: Termos apresentados por ordem decrescente de frequência.

Vemos que as palavras-chave associadas ao discurso associativo representam apenas 16% do total de frequências dos termos e um quarto do total de vocábulos integrados na análise (isto é, com mais de meia centena de repetições). As expressões mais frequentes incluem os termos “social” (rede/setor/área/ação/intervenção/segurança social), “política” (pública, social, de desenvolvimento, associativa) e “comunidade”. De resto, os diversos termos usados para identificar os públicos-alvo da atividade das organizações e que não constam da lista de termos selecionados (por não atingirem, individualmente, o número mínimo de referências) – beneficiários, destinatários, utilizadores, participantes, públicos-alvo, populações –, são globalmente relevantes no discurso das pessoas entrevistadas.

Ao contrário do que seria de esperar, **o background institucional não parece influenciar significativamente o uso de linguagem associativa**. Ou seja, ela não é substancialmente mais expressiva entre representantes do TS do que entre representantes de consultoras privadas⁷⁰ ou de universidades e entidades financiadoras (públicas). A exceção são os/as representantes das organizações de cúpula do terceiro setor, onde a defesa do jargão do setor é mais evidente.

A orientação científica da linguagem, por seu turno, concentra-se em grande medida na palavra-chave que deu tema à entrevista - “avaliação” -, e que é, de longe, a mais usada em contexto de entrevista (com 2002 referências, seguida de “financiamento” com um volume muito menor de referências - 451). Os restantes termos oriundos da cultura científica somam um peso de cerca 20% na contagem global da frequência de palavras-chave e um peso de pouco mais de um quarto no total de termos apurados para a análise. Ou seja, **são relativamente poucos os termos científicos popularizados nas organizações**. Além de “avaliação”, alguns dos termos mais frequentes, como “resultados”, “objetivos” e “impactos” são categorizados não apenas no domínio científico, mas também no gestor, já que se trata de termos com tradição forte em ambos os domínios, traduzindo, por um lado, um determinado enfoque no desenho da avaliação, e espelhando, por outro lado, uma abordagem de gestão orientada para os resultados. É exatamente ao

⁷⁰ Esta tendência geral é contrariada em relação a alguns termos, particularmente os termos usados para designar a população-alvo (beneficiários, destinatários, utilizadores, etc.).

nível da linguagem científica que as diferenças segundo características do *background* institucional das pessoas entrevistadas são mais salientes. **Claramente ela surge mais frequentemente no léxico de especialistas e profissionais de avaliação** associados a universidades e a consultoras.

Existe uma diversidade de termos oriundos da cultura científica, sobretudo presentes na teoria sobre avaliação (em especial sobre avaliação de programas) e que são relativamente comuns no discurso de pessoas com experiência na condução de estudos de avaliação e que se encontra debilmente difundida nas OTS (por exemplo, “pesquisa”, “critério”, “amostra”, “qualitativo”, “experimentação”). **Parece penetrar mais facilmente a linguagem das organizações a variedade de termos com origem na gestão**, os quais estão presentes numa série de debates no quadro de diversas tradições disciplinares (teoria organizacional, teoria da gestão, desenvolvimento internacional, etc.).

A maioria dos termos com maior número de ocorrências tem, pois, origens na cultura gestonária. Palavras como “financiamento”, “planeamento”, “gestão”, “estratégia”, “qualidade” e “certificação” são dominantes no discurso sobre avaliação nas OTS. Qualquer um deles é referenciado pelo menos quatro vezes mais (nalguns casos 10 vezes mais) do que termos importantes categorizados como pertencendo ao domínio associativo – “missão”, “cidadania”, “participativo”. Aliás, alguns termos caros ao associativismo e que surgem na literatura ligados ao debate sobre avaliação nas OTS, como “justiça”, “transparência”, “valores”, “confiança” ou “compromisso”, estão praticamente ausentes do discurso produzido pelas pessoas entrevistadas. O predomínio da linguagem com raízes na gestão dá suporte à ideia da crescente orientação instrumental do terceiro setor (Hwang e Powell, 2009; Stewart, 2014; Powell *et al.*, 2012) e aproximação ao perfil empresarial (Hvenmark, 2013; Maier *et al.*, 2014). **A forte influência do discurso gestonário parece assombrar o caráter “expressivo” das organizações** (isto é, de expressão de valores e ideais de serviço), **secundarizado nas narrativas sobre avaliação em favor do seu caráter “instrumental”** (material, objetivo).

As questões financeiras assumem grande centralidade no discurso das OTS sobre a avaliação em diversos contextos. Surgem por referência quer aos recursos para desenvolver as suas atividades, nomeadamente de avaliação (identificadas pelos termos

“financiamento”, “orçamento”, “receitas”, “recursos”), quer ao tipo de procedimentos, enfoques, conteúdos, requisitos e critérios de reporte e avaliação (identificadas pelo termo “financeiros”). Paralelamente, assume grande relevância no discurso uma variedade de termos e expressões ligados à gestão (estratégica) do desempenho – “gestão”, “planeamento”, “estratégia”, “objetivos”, “resultados”, “impacto” – e à garantia de qualidade – “qualidade”, “certificação”, “auditoria”, “melhoria” (sugestões/ações de melhoria, melhoria contínua/interna).

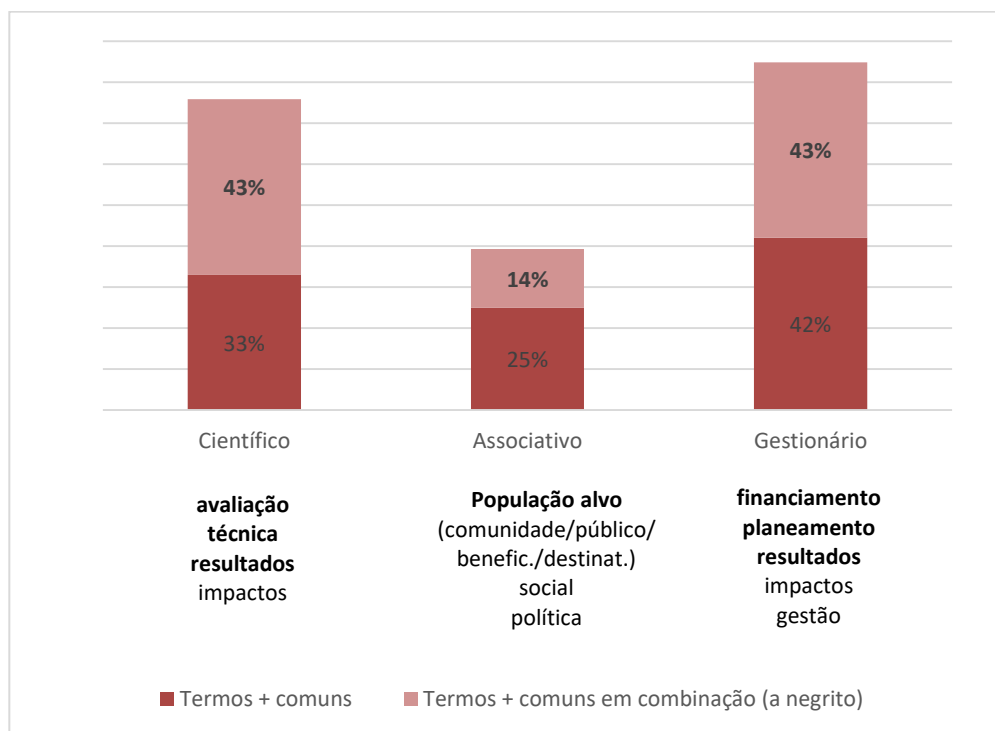
Apesar de a análise do tipo de vocabulário usado nas entrevistas produzir evidências de colonização pela ciência e pela gestão do debate sobre a avaliação nas OTS (Powell *et al.*, 2012), também observamos **sinais de combinação dos três tipos de linguagem**. Existe uma série de termos usados por grande parte das entidades (80%), em combinação com diversas linguagens, a que Powell *et al.* (2012) se referem como “interlinguagem”. Segundo os/as autores/as, a interlinguagem, sendo composta por elementos das linguagens originais, **constitui o discurso distintivo da avaliação no terceiro setor e torna possível a comunicação e a coordenação através das fronteiras sociais e linguísticas**.

Todavia, é relativamente restrito o grupo de palavras-chave usadas (individualmente) pela generalidade das organizações auscultadas, como se vê no gráfico abaixo. São 10 os termos usados por metade das entidades auscultadas, sendo que apenas 1 em cada 4 combinam todos os termos individualmente referenciados. Nesta linha, se consideramos apenas os vocábulos usados em combinação pela generalidade das entidades (pelo menos 80%) – a interlinguagem – o grupo de termos é reduzido para perto de metade. Segundo Galison (1997 *apud* Powell *et al.*, 2012), a interlinguagem designa pidgins⁷¹ e crioulos que emergem nos interstícios entre os domínios sociais, facilita a comunicação local através de fronteiras sociais e linguísticas e permite a coordenação da ação em todo o lugar, tempo e contexto. Os diversos termos que designam a população-

⁷¹ Também chamada de língua de contacto, é o nome atribuído a qualquer língua que é criada, normalmente de forma espontânea, a partir da mistura de outras línguas, e serve de meio de comunicação entre falantes de idiomas diferentes (Sebba, 1997).

alvo⁷² das organizações (comunidade, público, destinatários, beneficiários, utilizadores, participantes) são a marca mais visível da génese associativa da linguagem partilhada, claramente dominada pela orientação científica e gestonária, como se vê no gráfico abaixo.

Gráfico 2. Palavras-chave mais comuns e “interlinguagem”
(N= número de pessoas entrevistadas)



Fonte: Entrevistas semiestruturadas.

O objetivo de serviço à comunidade ou a grupos específicos de pessoas articula-se com preocupações técnicas em torno do planeamento, financiamento e apresentação de resultados. Embora a linguagem partilhada sobre avaliação nas organizações do terceiro setor estenda as fronteiras dos domínios da sociedade civil, investigação científica e gestão, o rácio aproximado de 3:1 (relação entre n.º de palavras-chave dos domínios

⁷² Considerando o peso global nas narrativas que têm as diversas palavras usadas para designar a população-alvo das organizações, que embora pertençam ao mesmo campo lexical (i.e. são relacionadas pelo sentido) não integram a mesma família de palavras (i.e. não estão relacionadas morfológica e etimologicamente), optei por agregá-las na análise da interlinguagem.

gestionário/científico e o n.º de termos do domínio associativo, que compõem a interlinguagem) **evidencia a colonização pelas linguagens científicas e gestionárias**⁷³.

O campo da avaliação do terceiro setor em Portugal está em mudança, e com ela ocorre também uma alteração ao nível do vocabulário. Ainda que expressões emergentes no campo da avaliação em Portugal como “retorno social do investimento” ou “análise contrafactual de impactos” ainda não tenham permeado o léxico da maioria das entidades do terceiro setor auscultadas, alguns termos como “planeamento”, “financiamento”, “resultados” ou “impactos” assumem grande centralidade na discussão sobre avaliação nestas organizações. O que esta análise torna visível é que com as crescentes pressões no sentido de as OTS reportarem sobre o seu trabalho, muitas das soluções adotadas são fortemente influenciadas por princípios de mercado. **A preocupação que daqui decorre é que a predominância do discurso gestionário (e científico) possa estar a excluir as vozes do terceiro setor e a promover sua orientação instrumental, perdendo de vista os seus valores constitutivos. Mas também não se pode dizer que a influência da cultura científica gestionária anule à partida a afirmação da cultura associativa.** Esse risco parece existir, mas também **há evidências de combinação e até reforço mútuo.** Vemos que em algumas organizações com processos de gestão profissionalizados ou com processos de avaliação concebidos e implementados profissionalmente, a forte utilização de linguagem com raízes na cultura gestionária e/ou científica articula-se com o uso de um conjunto relativamente diversificado de termos que refletem valores associativos. Por exemplo, entidades com processos de gestão e avaliação integrados no sistema Certificação da Qualidade dos Serviços Sociais (EQUASS) tendem a articular uma presença forte de termos como “gestão”, “qualidade”, “melhoria” e “clientes” com importantes termos com raízes na cultura associativa como “missão”, “princípios”, “visão”, “empowerment” e “envolvimento”. No caso de entidades com experiência na realização de estudos formais de avaliação (designadamente no âmbito de projetos ou respostas sociais específicas), vocábulos como “dados”, “indicadores”, “qualitativo”, “quantitativo” e “impactos”

⁷³ O único estudo identificado com um âmbito semelhante ao desta análise foi conduzido em contexto anglo-saxónico, e produziu resultados diferentes. Todavia, os resultados do trabalho da equipa de investigadores/as da Universidade de Stanford (Oelberger *et al.*, 2012; Kloos e Papi, 2014; Korff, Oberg e Powell, 2015) não são comparáveis aos deste trabalho devido a múltiplas diferenças nos critérios de seleção das fontes e da amostra, bem como do conjunto de palavras-chave mobilizado na análise.

compatibilizam-se com termos caros ao associativismo, como “mudanças”, “benefícios”, “comunidade” e “participativo”.

4.2. A subscrição efusiva da retórica da avaliação

O pano de fundo do discurso das organizações a propósito das adoção das práticas de avaliação é, em geral, o **da internalização do “mantra” da avaliação**. Esta adesão à ideia da desejabilidade da avaliação e da sua utilidade instrumental encontra-se enraizada nos **pressupostos de racionalidade** que têm regido muita da discussão teórica e metodológica sobre avaliação. Embora, como veremos, a adesão à ideia da relevância da avaliação não seja necessariamente proporcional à intensidade das práticas de avaliação em vigor nas organizações, ela marca a elocução das pessoas entrevistadas quando se lhes pergunta diretamente sobre o que motiva os exercícios de avaliação. É por isso comum que, num primeiro momento, eles sejam apresentados como processos impulsionados internamente (pelo pessoal técnico ou dirigente das organizações), com propósitos instrumentais de melhoria do desempenho, seja em termos de processos seja em termos de resultados.

Desde a constituição da X tivemos sempre, desde o princípio, a preocupação de avaliarmos tudo o que fazíamos. Por outro lado, (...) o nosso presidente (...) era administrador de dezenas de empresas. Uma pessoa (...) muito conhecedora de várias áreas. A nossa preocupação desde o início foi: em qualquer projeto que quiséssemos implementar, irmos procurar quem nos pudesse supervisionar em termos científicos e que nos ajudasse na avaliação. (...) A ideia sempre foi avaliar e assim fazer com qualidade. (Entrevista com dirigente de ONG - 32)

Esta análise imediata das motivações para a avaliação veiculadas no discurso de representantes de organizações de base do terceiro setor é coerente com as veiculadas no inquérito por questionário. Assim, pediu-se às entidades para selecionarem **uma de um conjunto** de afirmações sobre **impulsionadores e desígnios da avaliação**:

- Fazemos avaliação porque os financiadores exigem;
- Fazemos avaliação porque nos ajuda a promover-nos perante a comunidade local;
- Fazemos avaliação para nos ajudar a tomar decisões organizacionais ou programáticas;
- Fazemos avaliação para nos promovermos perante os financiadores e atrair recursos.

Cerca de três quartos das entidades identifica como impulsionador da avaliação o seu papel de suporte à tomada de decisão (como previsto pela teoria da escolha racional),

sendo muito menos aquelas que remetem para a primazia de fatores associados às pressões do ambiente, como sejam as exigências das entidades financiadoras (como prenunciado pela **teoria da agência**), ou a atração de recursos (como previsto pela **teoria da dependência**), selecionados por pouco mais de 15% das entidades.

Estudos empíricos conduzidos em Portugal (Mesquita, 2015) e noutros países (Hall *et al.*, 2003; Salamon, Geller e Mengel, 2010; MacIndoe e Barman, 2012; Melloni, 2014; Mitchell, 2014; Mitchell e Berlan, 2016), nos quais foi pedido às organizações que identificassem os fatores que motivaram a adoção de prática de avaliação/medição do desempenho, secundam estes resultados, apontando para a primazia dos fatores internos na decisão de avaliar. Tais resultados sugerem que, mesmo quando a introdução de práticas de medição do desempenho ou avaliação é **induzida pelo financiamento**, estas **tendem a ser normativamente internalizadas como parte da boa gestão**, prática e governação organizacional (Ellis e Gregory, 2008; Harlock, 2013).

De resto, **peso das motivações internas** de melhoria e procura de qualidade só tem relevância no discurso de dirigentes e técnicos/as das organizações de base do terceiro setor, sendo claramente **reduzido, ou mesmo abandonado, no discurso de outros atores relevantes** no domínio do TS e da avaliação, nomeadamente em organizações de cúpula e consultores/profissionais de avaliação. No discurso de alguns destes atores, as **pressões externas têm lugar preponderante nas práticas das OTS ao nível da avaliação**, consubstanciando não só pressupostos da teoria da agência e da dependência de recursos, mas também forças miméticas e normativas do ambiente.

Há uma série de desafios que vêm do exterior das organizações que as levam a seguir esse caminho. Mas é uma questão mais reativa do que ativa. A maior parte das OTS reagem a esses desafios, mais do que antecipá-los e criar instrumentos para estar preparadas quando essas mudanças se dão. (Entrevista com representante de organização de cúpula – 43)

Em todo o caso, a **grande generalidade das entidades auscultadas através de inquérito e de entrevista (90%) declara fazer pelo menos alguma monitorização e avaliação das suas atividades**. A percentagem residual de entidades inquiridas que assumem não fazer ou fazer muito pouca monitorização das suas atividades é composta sobretudo por coletividades. Esta disposição geral é coerente com os resultados de estudos realizados em território nacional e noutros países. Num inquérito conduzido em Portugal

abrangendo uma amostra de Cooperativas de Solidariedade Social (CERCI), a mesma percentagem de entidades afirmava avaliar regularmente os seus programas e projetos (Nicolau e Simaens, 2008). Num inquérito abrangendo 31 organizações a trabalhar no setor da Cooperação para o Desenvolvimento em Portugal (não apenas ONGD, mas outras entidades públicas e privadas), a percentagem de adesão é mais baixa (60%), embora tal discrepância possa ser explicada pelos termos mais estritos em que é colocada a questão nesse estudo (Mesquita, 2015)⁷⁴. Em dois inquéritos às OTS conduzidos em contexto dos EUA apurou-se que mais de 80% (entre 84% e 90%) das organizações realizaram algum tipo de monitorização ou avaliação no ano precedente (Carman, 2007, 2008; Hoefler, 2009; Salamon, Geller e Mengel, 2010; Innovation Network, 2010, 2012, 2016) apesar de apenas uma pequena parte dessas entidades contar com um membro da equipa dedicado à avaliação em tempo integral. Inquéritos conduzidos em contexto europeu produzem uma maior diversidade de resultados, ainda que alinhados com a tendência de forte adesão à avaliação (Ellis e Gregory, 2008; Ogain *et al.*, 2012; Melloni, 2014). **Todavia, quando nos reportamos a práticas concretas, as organizações comportam-se de forma algo dissonante.** Esta dissonância entre discursos e práticas é particularmente visível quando, em contexto de entrevista, se pergunta às organizações que peso tem a monitorização e a avaliação no total das atividades conduzidas. As respostas apontam para a secundarização do papel da avaliação relativamente às demais funções das organizações e do pouco tempo que lhe é dispensado:

Em termos de tempo dispensado a ela, face a toda a execução, acompanhamento, recrutamento... é reduzida. O tempo que é dispensado para a avaliação nunca é superior a 5, 10%. (Entrevista com direção executiva de escola profissional – 15)

Apesar da desejabilidade e relevância da avaliação nos discursos, a reflexão sobre as intervenções, sobre os respetivos processos e resultados, é claramente **um domínio não prioritário de ação das organizações**. Por outro lado, como foi referido no início deste capítulo, o conceito de avaliação que emerge dos discursos e práticas das pessoas das organizações é muito abrangente, incluindo desde práticas com algum nível de sistematicidade e formalidade, a práticas de carácter informal mas que servem o propósito

⁷⁴ Nesse estudo a questão é colocada nos seguintes termos: “A organização realiza/contrata avaliações com regularidade?” (Mesquita, 2015: 32).

de melhorar o desempenho. Além de descreverem atividades tradicionais de monitorização e avaliação de projetos, assentes em metodologias das ciências sociais, muitas pessoas responsáveis de OTS entrevistadas referiram também um conjunto amplo de atividades de reporte, gestão e qualidade. Ponderando as respostas dadas em entrevista, optou-se por incluir no inquérito um conjunto alargado de práticas de controlo, gestão e avaliação referenciadas nas entrevistas.

Tabela 4. Práticas de gestão, controlo e avaliação*

(N=575)	N	%
Reporte		
Relatórios anuais (contas/plano atividades)	527	91,7
Relatórios para entidades financiadoras sobre atividades de projetos	335	58,3
Relatórios para entidades financiadoras sobre execução financeira de projetos	315	54,8
Controlo da qualidade e auditoria		
Processos certificação/acreditação	147	25,6
Auditorias financeiras externas	143	24,9
Auditorias técnicas/visitas de acompanhamento externas	184	32,0
Avaliação		
Avaliações formais de projetos	234	40,7
Análises SROI – <i>Social Return on Investment</i>	60	10,4
Conceção de “modelos lógicos”	117	20,3
Avaliações formais do desempenho organizacional	142	24,7
Gestão estratégica do desempenho		
Avaliação de desempenho do pessoal ao serviço	210	36,5
Monitorização do cumprimento dos objetivos e metas organizacionais	300	52,2
Plano estratégico a 3 ou mais anos	196	34,1
Sistema de gestão “balanced scorecard”	63	11,0

Fonte: Inquérito por questionário.

Nota: * N.º e % de entidades que reportam envolver-se nas atividades regularmente.

Vemos que dentre os processos de gestão, controle e avaliação elencados, os que colhem maior adesão por parte das OTS são relativos ao reporte das atividades, exercícios de **cariz eminentemente descritivo**. Os processos de avaliação formais (à escala de projeto ou organizacional) são menos frequentes. Enquanto mais de metade das entidades afirma elaborar regularmente relatórios anuais de atividade e de execução física e financeira de projetos, são menos de metade as que dizem conduzir regularmente avaliações formais de projetos e de desempenho organizacional, e menos de um quarto aquelas que mobilizam

ferramentas de avaliação específicas, como sejam as análises SROI e os modelos lógicos. Estes dados sugerem que **embora as organizações estejam envolvidas numa série de atividades que se poderão associar a uma maior responsabilização ou à “boa gestão”, isso não significa que estejam a fazer avaliação, isto é, a desenvolver exercícios estruturados para apreciar o mérito ou o valor das suas intervenções.**

4.3. Os diferentes enquadramentos da avaliação nas organizações do terceiro setor: da gestão estratégica à gestão de projetos

A avaliação das organizações do terceiro setor pode ser conduzida a diversos níveis, nomeadamente à escala organizacional e à escala de um projeto individual. Uma das falácias que potenciam a deturpação dos resultados de avaliação apontada na literatura (e.g., Renz e Herman, 2016; Cutt e Murray, 2000) prende-se exatamente com o nível da avaliação, uma vez que a eficácia organizacional pode ser indevidamente inferida a partir de projetos específicos ou do desempenho individual.

Ligados aos níveis organizacionais em que as atividades de avaliação e medição do desempenho podem ocorrer estão os enquadramentos que poderão apresentar as práticas de monitorização e avaliação nos processos mais amplos das organizações. Isto não significa, todavia, que o envolvimento nestes processos (planos operacionais, planos estratégicos, certificação de produtos/serviços, desenvolvimento de projetos), que poderemos associar à prestação de contas ou à “boa gestão”, implique necessariamente a condução de avaliações – entendidas como apreciações sistemáticas sobre o desempenho. Em todo caso, **é em alguns destes processos que se situam as alavancas para o surgimento e para a institucionalização das práticas de avaliação nas organizações.**

4.3.1. Os projetos como enquadramento privilegiado dos exercícios formais de avaliação

A **participação em projetos** é, de resto, uma forte **impulsionadora da avaliação nas organizações estudadas**, sobretudo projetos ou programas suportados por fundos públicos (nacionais ou europeus), como é o caso das iniciativas cofinanciadas pelo FSE. A par da certificação, a experiência de implementação de projetos e programas é a mais

salientada no inquérito, na resposta aberta à questão sobre a existência de alguma iniciativa marcante para a organização do ponto de vista da integração de práticas de avaliação (referida por **mais de um terço das organizações** que destacaram alguma iniciativa). Outro indicador da importância deste enquadramento da avaliação nas OTS refere-se à prevalência com que as entidades que respondem ao inquérito elaboram relatórios sobre a execução física e/ou financeira dos projetos: 61% afirma fazê-lo com regularidade. Ou seja, tendo em conta que estes relatórios são de natureza obrigatória, podemos deduzir que mais de metade das entidades promove projetos numa base regular.

De resto, os **exercícios formais de avaliação**, envolvendo a recolha e análise sistemáticas de informação qualitativa e quantitativa sobre processos e resultados, **mais enformados pelo campo formal da pesquisa em avaliação e alinhados com o tipo de discurso científico, ocorrem sobretudo ao nível dos projetos**. A análise das entrevistas e dos documentos disponibilizados pelas OTS vai ao encontro da perspetiva que emerge do inquérito, na qual a adesão à avaliação formal à escala dos projetos se destaca face à da escala organizacional (vide tabela 4). Embora se possa questionar a qualidade dos exercícios de avaliação que em geral são suscitados pelos projetos financiados, encontramos, nos relatórios de avaliação dos projetos, as análises mais sofisticadas e exigentes (em termos técnicos) do desempenho das intervenções. Encontramos uma maior diversidade e riqueza de metodologias (qualitativas e quantitativas), um maior envolvimento das partes interessadas e um maior recurso a especialistas em avaliação, embora o recurso à avaliação externa (não sendo aqui tomado como sinónimo de auditoria externa) seja, desde o fim do PIC EQUAL, pouco frequente.

Dentre a diversidade de iniciativas cofinanciadas pelo FSE, o **Programa de Iniciativa Comunitária EQUAL** assumiu relevância particular no que toca à **integração de práticas estruturadas e aprofundadas de avaliação** de projetos nas OTS, pelo especial investimento na promoção e organização de práticas de avaliação dos projetos. Através de um dispositivo de avaliação obrigatório, que conjugava a autoavaliação (envolvendo todos os atores “pertinentes” diretamente implicados na execução do projeto) e a avaliação “independente” (da responsabilidade de um avaliador “externo” ou de um elemento “interno” mas com autonomia face à equipa que executa as atividades), procurava-se

“inscrever uma dinâmica avaliativa que permitisse informar a gestão em cada momento de como a intervenção estava a decorrer” (entrevista com técnico/a do gabinete de gestão EQUAL) e difundir uma cultura de avaliação e de responsabilidade coletiva (Figueirôa e Pimenta, 2005). O inquérito realizado no âmbito deste estudo atesta o papel relevante da EQUAL no domínio da avaliação, ainda que a **intensidade da participação neste programa seja menos expressiva que a do POPH: envolveu uma em cada 10 entidades inquiridas**. Nas entrevistas, a **EQUAL** é apontado por **todas** as organizações que nela participaram na qualidade de entidades interlocutoras (15, no total de 31 OTS) como a **iniciativa mais marcante do ponto do impulsionamento das práticas de avaliação ao nível dos projetos**, a uma distância de 10 anos do envolvimento nesta iniciativa. O testemunho que se segue é representativo do discurso produzido pelas entidades a este respeito:

Eu estou nesta organização desde que ela foi criada, há 17 anos. Eu acho q há 2 marcos muito importantes em termos de trabalho: um tem a ver com a EQUAL. (...) Foi assim que eu entendi a iniciativa, foi assim que foram preparadas as equipas, um bocadinho para responder à iniciativa, como um desafio do ponto de vista da metodologia, da formação, das competências, muito interessante. Ou se trabalhava a sério nessa altura ou então não valia a pena... (Entrevista com coordenação técnica de ADL – 26)

Esta aposta antevia o potencial de disseminação do modelo de avaliação preconizado pelo programa. Todavia, **as expetativas de disseminação das práticas inovadoras ao nível da avaliação conduzidas no âmbito da EQUAL foram, pelo menos em parte, logradas pela descontinuação do programa no período de programação seguinte**, como emerge na seguinte perspetiva colhida em contexto de entrevista:

O grande desafio da EQUAL, e que não foi cumprido, era criar essa dinâmica de avaliação. Era dar um passo mais na criação de uma dinâmica de avaliação em entidades na intervenção social. E isso foi alcançado até um certo nível: e tocou, e ficou nas pessoas e ficou em algumas práticas. Mas depois, a rutura e a interrupção desse modelo, é absolutamente contraproducente face a esse objetivo. (...) Um programa como o EQUAL tem que ter continuidade. A descontinuidade deste espaço limita em si mesmo o seu alcance. (Entrevista com técnico/a do gabinete de gestão EQUAL - 2)

A estratégia seguida no período de programação dos fundos estruturais seguinte (2007-20013) foi a de integração da inovação e de outras lógicas experimentadas em programas como o EQUAL (trabalho em parceria, avaliação, transnacionalidade, igualdade de género, ...) nos programas *mainstream*, mediante a bonificação de projetos que contemplassem essas dimensões. **Todavia, pelo menos no que toca à avaliação, a lógica do PIC EQUAL não foi retomada no quadro comunitário que lhe sucedeu:**

Esses programas [experimentais] desenvolveram-se paralelamente e os seus resultados nunca foram tomados a sério ou incorporados nas práticas correntes, quer mesmo aquelas que eram financiadas através de Fundos Estruturais e com cofinanciamento nacional, que é o caso do POPH dentro do QCA, ou até de programas de natureza estritamente nacional que pudessem beneficiar desses resultados. (Entrevista com especialista em avaliação - 38)

Um dos Programas que, pela proximidade da agenda temática e de alguns dos objetivos que animam, poderia integrar alguns resultados e aprendizagens carreados pela iniciativa EQUAL seria o Programa Operacional Potencial Humanos (POPH). A abrangência e amplitude do programa é atestada no estudo, já que perto de metade das organizações inquiridas estava a promover ou já tinha promovido pelo menos um projeto POPH, sendo em média cinco o número de projetos em que cada uma diz estar ou ter estado envolvida. A maior fatia do investimento do POPH é em ações formativas (70% dos recursos destinados à qualificação de nível secundário), por isso não é de estranhar que, refletindo processos de isomorfismo institucional, a maior parte dos projetos apoiados destas organizações tivesse as ações de formação como principal atividade.

Ao contrário do PIC EQUAL, no enquadramento do POPH **não existe obrigatoriedade de apresentação de relatório de avaliação dos projetos**. As entidades são apenas **obrigadas ao reporte de informação numérica** de execução física e financeira dos projetos, através de submissão ao SIIFSE (Sistema Integrado de Informação), com base nos indicadores de acompanhamento e desempenho definidos para o programa.

Fazemos um relatório do projeto, no caso do POPH, embora não seja obrigatório no projeto, porque eles próprios, a plataforma, consideram que ao inserir a informação e os dados físicos, para eles basta isso. Mas nós fazemos para o relatório. (Entrevista com técnico de ONG – 14)

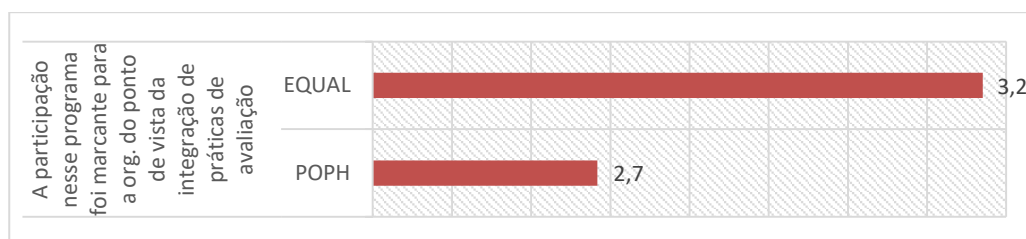
No documento de programação do POPH (POPH, 2014) as indicações sobre a avaliação no âmbito dos projetos estão ausentes, ainda que se estabeleçam para o programa uma série de indicadores de acompanhamento e de desempenho que alimentarão o SIIFSE e em cuja recolha estarão envolvidos os promotores dos projetos, com o objetivo de “proporcionar ao sistema de gestão e acompanhamento do Programa uma informação regular sobre o estado da execução” e “aferir sobre o desempenho do PO na concretização dos seus objetivos”. A ausência de orientações ou requisitos para avaliação no âmbito dos projetos é só parcialmente colmatada na regulamentação específica de cada tipologia de intervenção. Apesar de na generalidade dos formulários de candidatura a

projetos as organizações promotoras serem instadas a identificar os “mecanismos estruturados de avaliação da qualidade e dos resultados do projeto”, em boa parte das tipologias os mecanismos de avaliação propostos não são considerados nos critérios de seleção das candidaturas nem constam das respetivas grelhas de análise. Além disso, como reconhece responsável da gestão do programa entrevistado, não houve uma verificação da real concretização do que foi expresso em sede de candidatura:

As instituições quando se candidatavam, que tinham que identificar quais as metodologias ou mecanismos (...) para avaliação. Mas depois reconheço que a monitorização ou verificação do cumprimento desse requisito ficou.... Não é, nem tem sido verificado no final. (Entrevista com responsável órgãos de gestão POPH - 3)

O posicionamento das entidades relativamente ao impacto da participação nos dois programas no que toca à avaliação é revelador da maior relevância da EQUAL face ao POPH: são significativamente mais as entidades que tendem a concordar que a participação na EQUAL foi marcante para a organização ao nível da integração das práticas de avaliação⁷⁵.

Gráfico 3. Perceções sobre impacto do POPH e EQUAL na organização*



Fonte: Inquérito por questionário.

Nota: *Valores médios: 1 = Discordo totalmente; 2 = Discordo; 3 = Concordo; 4 = Concordo totalmente

Ao contrário do observado no caso do EQUAL, nas entrevistas a experiência de implementação de projetos no âmbito do POPH raramente é apontada como particularmente marcante no que toca à internalização da avaliação.

A ideia da desvalorização da avaliação ao nível dos projetos nos últimos anos (a partir do fim da EQUAL) não se confina ao POPH, mas estende-se à generalidade da intervenção cofinanciada pelo Estado português de que são beneficiárias OTS. A percepção

⁷⁵ Para testar se existiam diferenças entre POPH e EQUAL nos valores médios de adesão à afirmação foi realizado o teste de Wilcoxon para uma mediana. Concluimos, com uma probabilidade de erro de 5%, que os valores médios de adesão a ambas as afirmações são significativamente diferentes no POPH e no EQUAL ($p=0,003$).

dominante entre as pessoas envolvidas na gestão, implementação e avaliação de programas comunitários, especialistas e profissionais de avaliação e de representantes de OTS e respetivas organizações representativas é de que, a partir do último quadro comunitário, tem havido um **desincentivo à avaliação ao nível dos projetos, em favor das lógicas burocráticas de reporte descritivo sobre as despesas e realizações**. O excerto que se segue ilustra esta perspetiva:

Houve um desinvestimento nesta área da avaliação nos projetos. Programas como o EQUAL puseram em cima da mesa a questão da avaliação como uma questão importante. Se, de repente, a mensagem que o Estado passa o dinheiro para as autarquias e instituições (...) sem exigência de avaliação, então eu que estou sentado nas organizações vou dizer: então porque é que eu vou investir numa coisa se nem o Estado valoriza". (Entrevista com especialista/profissional de avaliação – 37)

A este propósito, além dos projetos candidatados diretamente ao POPH pelas organizações no âmbito das diversas tipologias, a pesquisa empírica fez sobressair **dois programas governamentais que, beneficiando do cofinanciamento do FSE no âmbito do POPH, são promovidos por organismos públicos, sendo depois executados mediante o desenvolvimento de projetos liderados por OTS e outras entidades**: Os Contratos Locais de Desenvolvimento Social⁷⁶ e o Programa Escolhas⁷⁷. Ainda que estas iniciativas respondam aos requisitos do financiamento comunitário, a que respondem as entidades públicas que acedem diretamente aos apoios, a sua concretização em projetos obedece a uma série de orientações estabelecidas por aquelas entidades. **Algumas dessas orientações respeitam exatamente às atividades de monitorização e avaliação, que vêm acrescentar às dos fundos estruturais**, em particular, ao nível da autoavaliação, no caso dos CLDS e “Escolhas”, e ao nível da avaliação sumativa por parte da equipa técnica do Programa, apenas neste último⁷⁸.

⁷⁶ Os Contratos Locais de Desenvolvimento Social (CLDS) são promovidos pelo Instituto de Segurança Social e têm por finalidade promover a inclusão social dos cidadãos, de forma multissetorial e integrada, através de ações a executar em parceria, para combater a pobreza persistente e a exclusão social em territórios deprimidos. (<http://www4.seg-social.pt/contratos-locais-de-desenvolvimento-social-clds>)

⁷⁷ O Escolhas é um programa governamental de âmbito nacional, criado em 2001, promovido pela Presidência do Conselho de Ministros e integrado no Alto Comissariado para a Imigração e Diálogo Intercultural – ACIDI, IP, cuja missão é promover a inclusão social de crianças e jovens de contextos socioeconómicos vulneráveis, visando a igualdade de oportunidades e o reforço da coesão social. (<http://www.programaescolhas.pt/apresentacao>)

⁷⁸ Este programa contempla um processo de autoavaliação, cujo modelo de relatório é fornecido pelo Programa, complementado pela utilização obrigatória de uma ferramenta informática de avaliação

Independentemente do modelo de avaliação preconizado, estes dois casos vêm temperar a percepção do refreamento do papel da avaliação de projetos apoiados por fundos públicos nos últimos anos. Não obstante, apesar de o legado da avaliação do PIC EQUAL, especialmente da autoavaliação, não ter sido inteiramente negligenciado pela gestão das iniciativas que lhe sucederam, certo é que em nenhuma delas foi tão ambiciosa quer em termos de requisitos quer em termos de recursos para a avaliação, nomeadamente para a avaliação externa. O forte investimento na avaliação dinamizou o mercado, com o aumento da procura e da oferta de prestação de serviços de avaliação. **Uma vez que a avaliação independente, nomeadamente externa, deixou de ser exigida ou encorajada no quadro comunitário seguinte, é razoável a dedução de que houve uma retração do mercado de avaliação ao nível dos projetos,** conforme preconizado na alínea b) da terceira hipótese do estudo.

4.3.2. A prevalência da medição do desempenho na avaliação à escala organizacional

A avaliação à **escala organizacional** tende a assumir feições próximas das **práticas da gestão empresarial**, sobretudo associadas à gestão estratégica e do desempenho (verificação do cumprimento dos objetivos e metas organizacionais e avaliação de desempenho do pessoal ao serviço). **Nestes exercícios à escala organizacional** as análises mais compreensivas e contextuais que têm lugar à escala dos projetos tendem a ceder lugar a uma **lógica mais alinhada com os pressupostos da medição do desempenho** – entendida como processo de reporte de “características observáveis e quantificáveis” de projetos e práticas do terceiro setor (Volkman, 1999: 27) – em maior articulação com sistemas de gestão mais amplos, como a gestão estratégica e da qualidade. A forte adesão à monitorização do cumprimento de objetivos e metas organizacionais e à avaliação do

online também fornecida pelo programa. Estes relatórios terão obrigatoriamente que espelhar não apenas a perspetiva do consórcio mas também das pessoas destinatárias, mediante a realização de assembleias de jovens. Os CLDS também disponibilizam guião para o processo de autoavaliação (semelhante ao usado na EQUAL) mas não prescrevem o envolvimento dos/as destinatários/as no exercício. Também não preveem avaliação interna, como no caso do Programa Escolhas. Aqui as conclusões da avaliação interna sumativa (traduzidas numa classificação) podem induzir alterações quer na condução dos trabalhos dos projetos, quer na continuidade do próprio financiamento.

desempenho individual – que envolvem com regularidade, respetivamente, mais de metade e de um terço das entidades inquiridas – é indicadora desta disposição. Em muitos casos, os sistemas de medição do desempenho são definidos e guiados por processos de planeamento estratégico (ou de mais curto prazo) ou de gestão da qualidade, sendo as ferramentas específicas de medição, como o *Balanced Scorecard*, pouco frequentes. **Esta lógica tem permeado também a avaliação conduzida no âmbito dos projetos**, nomeadamente do POPH, cuja informação requerida se destina a alimentar indicadores quantificados de desempenho do programa.

O enquadramento mais comumente enunciado nas organizações auscultadas para os exercícios avaliativos à escala organizacional é o dos planos de ação, mais concretamente os **planos anuais de atividades**, a que estão **legalmente obrigadas** as organizações do terceiro setor em Portugal. Em mais de 90% das entidades inquiridas e em todas as 31 entrevistadas este é pelo menos um dos enquadramentos das atividades de monitorização e avaliação conduzidas. A prevalência do planeamento de curto prazo, através planos de atividades anuais cuja execução é reportada através de relatórios de atividades, é salientada em análises anteriores (Ferreira *et al.*, 2008; Ruela, 2013; Parente *et al.*, 2014). De resto, o relatório que decorre do planeamento anual acaba por constituir **o reduto para alguma sistematização ou formalização das práticas de monitorização das pequenas organizações de carácter expressivo e voluntário**, como é o caso das coletividades⁷⁹. Mas, mesmo em algumas organizações que se afastam do perfil voluntário, os relatórios de atividades anuais constituem o único produto da reflexão sobre o desempenho. Parece ser sobretudo o caso de algumas IPSS entrevistadas, cuja atividade se concentra na prestação de serviços sociais, de média dimensão financeira, dependentes de financiamento da segurança social (mais de 70% do orçamento global), que não têm implementados sistemas de gestão da qualidade e/ou processos de planeamento estratégico.

⁷⁹ Dentre o leque diverso de processos de gestão, controlo e avaliação listados no inquérito às OTS (vide tabela 4), a elaboração de planos/relatórios anuais de atividade são os únicos processos em que a adesão das coletividades não é estatisticamente inferior à adesão de outros tipos de organizações (IPSS, ONG, ADL e outras) (*p-value* > 0,005) e em que a distância entre as pequenas organizações e as restantes não é tão acentuada.

Este tipo de planeamento, operacional, de curto prazo, é diferenciado, na literatura, do **planeamento estratégico** (e.g., Courtney, 2002; Drucker e Rangan, 2008), definido a mais longo prazo (3 ou mais anos), e que constitui "um esforço disciplinado para originar decisões fundamentais que deem forma e guiem uma organização" (Bryson, 2011), em direção à ideia de gestão estratégica, enquanto processo de mudança estratégica. O planeamento estratégico, que se propõe promover uma gestão eficaz priorizando metas num quadro de recursos limitados (Hwang e Powell, 2009: 272), enquadra-se no processo de gestão estratégica "através do qual as organizações determinam a sua missão, valores, visão, objetivos, estratégias, programas de ação e metas de desempenho desejadas (...) envolvendo várias técnicas e ferramentas para garantir uma formulação cuidada, uma implementação eficaz e eficiente, e a avaliação" (Anheier, 2005:259). **Uma das fases/dimensões essenciais do planeamento estratégico é exatamente monitorização e avaliação da implementação da estratégia formulada, mediante a apreciação sobre o alcance dos objetivos e metas estabelecidos**, e é por isso que ele aparece no discurso das pessoas entrevistadas como um enquadramento importante dos processos de reflexão interna. O uso do planeamento estratégico nas OTS surge associado à redefinição das missões e à exploração de estratégias para a geração de receitas, captação de recursos e relações públicas (Hwang e Powell, 2009). Análises recentes em contexto português apontam para um investimento crescente no planeamento estratégico, ainda que o planeamento anual, correspondendo a uma norma estatutária, continue a ser preponderante⁸⁰.

No presente estudo, **embora este tipo de planeamento a médio/longo prazo assuma grande relevância na acomodação das práticas de avaliação nas organizações que o desenvolvem (particularmente das práticas de nível organizacional), ele ainda não se encontra generalizado (envolve menos de um terço das entidades inquiridas e menos de**

⁸⁰ Um estudo dinamizado pelo CIRIEC, em 2008, no qual foi inquirida uma amostra de Cooperativas de Solidariedade Social (51), dava conta da adesão perto de 40% das entidades a processos formais de planeamento estratégico (Nicolau e Simaens, 2008). Estudos mais recentes apontam para o crescimento do envolvimento nestes processos. Em 61% das 153 OTS inquiridas no âmbito do estudo de Raquel Campos (2015), verifica-se a existência de processos de planeamento estratégico. Na linha desta tendência, no estudo coordenado por Cristina Parente (2014), verificaram-se práticas de planeamento estratégico em todas as 7 OTS estudadas em profundidade, ainda que no universo das OTS inquiridas a prática seja bem menos frequente – envolvendo menos de um quarto das organizações.

um quarto das entrevistadas)⁸¹. Ao contrário do planeamento de curto prazo, mais transversal aos diversos tipos de organizações, o planeamento estratégico é um processo mais seletivo em termos de estrutura organizacional⁸², quer ao nível do tipo de organização, perfil da atividade e do financiamento, dimensão financeira e antiguidade. Com efeito, tende a excluir sobretudo coletividades, organizações de pequena dimensão financeira (até 99.999€ receitas/ano), dependentes de receitas próprias, com funções de expressão e com ciclos de vida maiores (51 ou + anos), prevalecendo em organizações de maior dimensão financeira, antiguidade intermédia e dependentes de donativos ou do financiamento de projetos.

Um outro enquadramento comum das atividades de monitorização e avaliação conduzidos à escala organizacional, e o onde a **lógica da medição é prevaiente**, é a **certificação de produtos/serviços**. Trata-se de instrumentos formalizados, destinados a coordenar e a melhorar as operações de gestão (do trabalho, dos produtos, dos processos e do pessoal) através de procedimentos de resolução de problemas, sustentados numa nova forma de racionalidade de carácter processual (Cruz, pais e Parente, 2014: 132)⁸³. Para

⁸¹ Os valores mais baixos de adesão neste estudo em comparação com estudos anteriores conduzidos em território nacional serão em boa parte devedores das diferenças na composição da amostra. Enquanto na presente análise o peso das coletividades de cultura, desporto e recreio é considerável, nas análises anteriores este tipo de entidades é excluído da amostra.

⁸² Para avaliar se a adesão ao planeamento estratégico varia de forma estatisticamente significativa consoante as características da estrutura organizacional – dimensão financeira, antiguidade, principal fonte de receitas, tipo de entidade, funções principais – recorreu-se ao teste não paramétrico de Kruskal-Wallis, seguido da comparação múltipla da média das ordens, como descrito em Maroco (2014). Usou-se uma probabilidade de erro de tipo I (α) de 0,05. Verificou-se que, para todos os fatores considerados, verificam-se diferenças estatisticamente significativas (p -value=0,000).

⁸³ De acordo com listagens extraídas do *site* da Direção Geral do Emprego e das Relações de Trabalho (DGERT)⁸³ com informação relativa às entidades acreditadas/certificadas para o **domínio da formação**, existiam em 2014 cerca de **582 OTS acreditadas/certificadas**. Listagens disponibilizadas no [site - http://certifica.dgert.msess.pt/pesquisa-de-entidades-formadoras-certificadas-pela-dgert.aspx](http://certifica.dgert.msess.pt/pesquisa-de-entidades-formadoras-certificadas-pela-dgert.aspx) (dados extraídos em Abril de 2014). A seleção das entidades do terceiro setor foi feita a partir da informação sobre a demonização social.

O número de entidades com certificação na **área da qualidade** é inferior: **cerca de 381 OTS** – 230 pela norma ISO 9001⁸³, 116 pelo Modelo da Qualidade Europeu para os Serviços Sociais - EQUASS⁸³, e 35 pelo Modelo de Avaliação da Qualidade das Respostas Sociais – ISS (a que correspondem 68 respostas sociais certificadas).

- O número de entidades com certificação **ISO 9001:2008** foi apurado a partir da Base de Dados Nacional de Sistemas de Gestão Certificados. As listagens, extraídas em Abril de 2014 (em http://www.ipac.pt/pesquisa/pesq_empcertif.asp). A seleção das entidades do terceiro setor foi feita a partir da informação sobre a demonização social e âmbito da atividade. As normas da família ISO 9000 foram

além de acederem a um selo certificador da qualidade dos produtos e serviços prestados, as normas implicam também uma série de novos dispositivos que acrescem aos tradicionais instrumentos de gestão (Cruz, Pais e Parente, 2014), designadamente ao nível da monitorização e medição do desempenho: **dispositivos de avaliação de desempenho de colaboradores/formadores, avaliação da satisfação de stakeholders internos (colaboradores/as) e externos (clientes, parceiros, comunidade, financiadores, fornecedores), avaliação do impacto da formação e medição do desempenho (face a objetivos definidos)**. Embora alguns contemplem instrumentos de auscultação de diversos grupos de grupos de stakeholders, estes modelos, estruturados em torno de princípios da qualidade, são caracterizados por uma eminente orientação para os interesses e preferências do cliente – e não uma orientação para a ativação de um conjunto alargado de partes interessadas, como no caso da “avaliação de quarta geração” (Guba e Lincoln, 1989), sendo que a qualidade é medida principalmente pela avaliação de utilidade por parte dos clientes. Por esta razão, todos os processos e estruturas são orientados para a satisfação das suas necessidades⁸⁴.

desenvolvidas para apoiar as organizações, de qualquer tipo e dimensão, na implementação e gestão de sistemas da qualidade eficazes.

- O número entidades certificadas pelo **modelo EQUASS** foi apurado a partir da extração em Abril de 2014 das bases de dados das Entidades Certificadas EQUASS Assurance/EQUASS Excellence (em <http://www.equass.be/equass/index.php/equass-certification-system/our-certified-organisations/list-fabrik>, e http://www.equass.be/equass/index.php/equass-certification-system/our-certified-organisations/list-fabrik-exc?group_by=database_excellence_Country). Este modelo constitui uma iniciativa da EPR - *European Platform for Rehabilitation*, e é estruturada nos princípios da qualidade da *Common Quality Framework for Social Services of General Interest* e na metodologia da Auto-Avaliação.

- Por fim, a informação sobre entidades certificadas (cada entidade pode ter mais do que uma resposta social certificada) pelo **Modelo de Avaliação da Qualidade das Respostas Sociais – ISS**, foi enviada (entre Março e Abril de 2014), mediante requerimento, pelas entidades acreditadas para a certificação das respostas sociais, de acordo com os Modelos de Avaliação da Qualidade da Segurança Social (SGS ICS, EIC e APCER). Trata-se de um referencial normativo que se baseia nos princípios de gestão da qualidade e onde são estabelecidos os requisitos necessários à implementação do Sistema de Gestão da Qualidade dos serviços prestados pelas Respostas Sociais. Este modelo tem por referência a norma NP EN ISO 9001:2000 – Sistemas de Gestão da Qualidade e os Requisitos do Modelo de Excelência da European Foundation for Quality Management (EFQM).

⁸⁴ O lançamento, em 2003, do Programa de Cooperação para o Desenvolvimento da Qualidade e Segurança das Respostas Sociais, e subsequente edição dos primeiros Modelos de Avaliação da Qualidade das Respostas Sociais e das Recomendações Técnicas para Equipamentos Sociais, em 2007, constituem marcos decisivos da aposta na qualificação das organizações prestadoras de serviços sociais. Os programas de financiamento que integraram o III QCA - Quadro Comunitário de Apoio (2000-2006) e o QREN - Quadro de Referência Estratégica Nacional (2007-2013) potenciaram a implementação daquele programa já que

Cada um dos instrumentos referidos no Capítulo 1 (monitorização, avaliação, controlo, auditoria, medição do desempenho, etc.) é um conceito de direito próprio que é usado pela gestão empresarial e desenvolvimento da qualidade. Os sistemas de gestão da qualidade em conformidade com os modelos International Organization for Standardization (ISO), European Quality in Social Services (EQUASS) ou da Qualidade das Respostas Sociais (Instituto de Segurança Social) são modelos que, sendo mais abrangentes, enquadram estes instrumentos de gestão. Efetivamente, o presente estudo evidencia o papel de relevo que os sistemas de qualidade assumem no enquadramento e impulsionamento das práticas de avaliação de nível organizacional, mais especificamente das práticas de monitorização e medição do desempenho, nas OTS em Portugal. **Tal influência é bem reveladora dos efeitos de “isomorfismo institucional”, particularmente de tipo normativo, associados aos processos de acreditação/certificação.**

Quando falamos em **monitorização e avaliação com abrangência organizacional, os processos de gestão da qualidade** encontram-se, de facto, entre os **principais catalisadores, constituindo marcos relevantes na evolução da institucionalização destas práticas nas organizações estudadas.** No inquérito, das 131 pessoas que destacaram uma iniciativa/projeto em particular como tendo marcado a organização do ponto de vista da integração de práticas de avaliação, 33 (cerca de um quarto) referiram-se aos processos de certificação/acreditação. Os testemunhos abaixo, colhidos nas entrevistas, são elucidativos desse papel:

O EQUASS foi decisivo para a integração destas práticas, não foi só um contributo. Se não fosse o EQUASS muitos dos procedimentos que foram adotados, a entidade não tinha capacidade de os implementar. (Entrevista com coordenação técnica de IPSS – 13)

No nosso caso, todo este processo de avaliação nasceu de uma cultura interna instalada que decorreu do facto de sermos a primeira [] com certificação ISO 9001. Portanto, isso aconteceu em início deste milénio, de tal forma que a partir daí estas práticas foram endogeneizadas. (Entrevista com dirigente/técnico de ONG - 23)

Com efeito, **mais de um quarto das entidades inquiridas e de um terço das organizações entrevistadas estão envolvidas em processos de certificação,** sendo os mais

possibilitaram o desenvolvimento de projetos com o objetivo de fomento das práticas de qualificação destas organizações.

frequentes a acreditação/certificação da formação DGERT e a certificação da qualidade de acordo com os modelos ISO 9001:2008 e EQUASS⁸⁵. A certificação de acordo com o Modelo de Avaliação da Qualidade das Respostas Sociais do Instituto de Segurança Social é menos frequente (surgiu em apenas 1 das 31 organizações entrevistadas), embora um maior número de entidades IPSS, especificamente, mencione a intenção de obter esta certificação e/ou afirme estar a implementar já práticas de qualidade utilizando como referência os instrumentos de gestão da qualidade do ISS, ainda que sem certificação externa. Algumas organizações em estudo projetam (ou em algum momento já projetaram) certificar-se, mas (ainda) não o fizeram por falta de tempo, de competências ou de recursos.

O envolvimento em processo de certificação é particularmente sensível às características da estrutura organizacional⁸⁶, quando comparado com os processos de planeamento operacional, correspondendo a um perfil de organização mais restrito. À semelhança do constatado em análises empíricas anteriores (Cruz, Pais e Parente, 2014; Franco, 2015), no caso da amostra inquirida neste trabalho a implementação de sistemas de gestão da qualidade e de certificação tem tido maior incidência nas IPSS do que nas Não IPSS. As IPSS e as ADL (muitas delas com estatuto de IPSS), com antiguidade intermédia, de grande dimensão financeira, exclusivamente prestadoras de serviços e dependentes de financiamento público, compõem o perfil das organizações certificadas no nosso país. Coerentemente com o constatado na análise de Cruz, Pais e Parente (2014), são as organizações mais antigas (com mais de 50 anos) e as mais jovens (até 20 anos) que revelam menores práticas de certificação. **A implementação do processo de certificação parece exigir alguma maturidade organizacional, pelo esforço que representa em termos de competências e recursos materiais e humanos,** “estando eventualmente as OTS mais

⁸⁵ Estudos anteriores conduzidos em território nacional apontam para resultados semelhantes. No inquérito conduzido no enquadramento do estudo coordenado por Cristina Parente (2014) 26% das 87 OTS inquiridas afirmava deter alguma certificação de produtos/serviços, sendo a conferida pela DGERT a mais frequente, logo seguida pela ISO 9001. O estudo conduzido por Raquel Franco (2015) conduziu a resultados mais expressivos: a maioria (63%) das organizações com estatuto de IPSS inquiridas já havia implementado ou estava a implementar um sistema de gestão da qualidade (face a 32% das não IPSS, num total de 153 entidades inquiridas).

⁸⁶ O recurso ao teste não paramétrico de Kruskal-Wallis permitiu concluir que, para todos os fatores considerados – dimensão financeira, antiguidade, principal fonte de receitas, tipo de entidade, funções principais – existem diferenças estatisticamente significativas ($p\text{-value}=0,000$).

enraizadas menos predispostas para empreender tais investimentos e as mais recentes algo carentes desta possibilidade” (Cruz *et al.*, 2014: 143). As entrevistas acolhem estas tendências, sobretudo a respeito das organizações jovens, algumas delas motivadas mas com poucas condições financeiras e humanas para empreender estes processos. Por outro lado, a **prestação de serviços diretos** está particularmente adaptada ao âmbito das certificações, percebidas como menos vocacionadas para a certificação de atividades de expressão. Sendo as organizações prestadoras de serviços sociais tendencialmente de maior dimensão e com uma forte componente de financiamento público (nomeadamente no âmbito dos acordos de cooperação com o Estado), estarão mais expostas às pressões das entidades tutelares no sentido de qualificarem a sua intervenção, além de que terão mais recursos para o fazer.

4.3.3. A centralidade da conformidade com normas e procedimentos nos processos de monitoria externa

Os processos de **monitoria externa**, amiúde nomeados enquanto exercícios de avaliação, são práticas correntes nas organizações. Os resultados do inquérito são reveladores. As auditorias técnicas e/ou financeiras da iniciativa das entidades financiadoras ou tutelares têm lugar regular ou ocasionalmente em perto de 60% a 70% das organizações, respetivamente. Este tipo de iniciativa ocorre numa diversidade de situações, desde processos de certificação da qualidade e da formação, até visitas de acompanhamento técnico da Segurança Social (no âmbito da prestação de respostas sociais) e auditorias financeiras e técnicas no âmbito de projetos comunitários ou internacionais. **Trata-se, pois, de exercícios que podem ter objetivos muito diversos embora, em geral, partilhem a orientação administrativa e normativa, e o enfoque nos processos de implementação.** Tendem a partilhar também a **orientação tradicional dos processos de auditoria**, centrando-se na validade material dos elementos objeto de auditoria e a sua conformidade global com determinados princípios regras e normas, e a **afastar-se das versões mais prospetivas da auditoria** como seja a auditoria de desempenho, e que são concetualmente mais próximas da avaliação.

No caso da **certificação**, as auditorias incidem sobre a verificação do cumprimento dos requisitos de certificação. Para a certificação da qualidade segundo os modelos ISO,

EQUASS e Sistema de Qualificação das Respostas Sociais (ISS), a metodologia de avaliação prevista tem por base a realização de **auditorias efetuadas por uma entidade externa reconhecida no âmbito do Sistema Português da Qualidade** e assenta sobretudo na revisão de documentos e realização de entrevistas⁸⁷. Similarmente, as práticas das entidades formadoras certificadas são também objeto de avaliação regular através de auditorias asseguradas pela DGERT, com recurso aos seus quadros ou a auditores/as externos/as.

Por seu turno, as **visitas de acompanhamento técnico** por parte dos serviços da Segurança Social, que podem ter lugar em estabelecimentos de apoio social (IPSS e outras entidades privadas com ou sem contratos de cooperação com o Estado), obedecem a uma lógica de auditoria ainda que com designação distinta. Os motivos das ações de acompanhamento podem ser diversos, podendo resultar de programação de acompanhamento definida, de denúncia/reclamação, de pedido de apoio financeiro ou de revisão de acordo de cooperação, entre outros. As visitas incidem sobre a verificação do cumprimento de determinados procedimentos, dependendo do âmbito da ação de acompanhamento (cumprimento de acordos de cooperação, processos de qualificação de resposta social, avaliação dos espaços, recursos humanos, etc.). A entrevista que se segue sintetiza o enfoque deste tipo de iniciativas das entidades financiadoras/tutelares:

As auditorias dos estabelecimentos são relacionadas com o funcionamento dos estabelecimentos e com as condições de segurança, higiene, qualidade do ar, nível de ruído (...). Depois há auditorias da tutela, no ano passado tivemos uma auditoria do ministério de trabalho e da Segurança Social, que foi à fundação toda e durou 4 meses. Tiveram aqui 3 técnicos durante todo esse tempo. Foi uma coisa grande, a todos os aspetos da fundação. (Entrevista com coordenação técnica de IPSS – 5)

No caso das **auditorias no âmbito de projetos, a amplitude de lógicas é maior**, variando consoante as entidades financiadoras. Por exemplo, organismos internacionais como a Comissão Europeia ou o Banco Mundial (particularmente no âmbito do apoio ao desenvolvimento) são identificados como financiadores particularmente exigentes e exaustivos ao nível dos processos de auditoria financeira e, em alguns casos, técnica. É

⁸⁷ Além das auditorias externas para aceder ou renovar a certificação, os modelos estimulam auditorias internas (frequentemente conduzidas por auditores/as externos/as) em intervalos planeados para determinar se o sistema de gestão da qualidade está conforme com as disposições planeadas e requisitos das normas e se estão a ser implementadas com eficácia.

exatamente no âmbito de projetos com financiamento de agências internacionais que encontramos processos de monitoria externa mais afinizados com a **auditoria de desempenho, apresentando maiores similitudes com algumas formas de avaliação**, designadamente em termos de fontes de informação, incidência e profundidade das análises. O caso da seguinte ONGD é ilustrativo:

E essas avaliações [no âmbito da cooperação] a que eles [UE] chamam de monitoria (...) são interessantes, muito pesadas também, muito duras, em termos de acompanhamento, resposta, exigentes, duras neste sentido. Porque o monitor tem uma margem de intervenção muito grande. Não é a produção de um documento inócuo, pode introduzir até alterações. (...) São pessoas que vêm de fora e que num mês ou 2 vão ter que conhecer 20 projetos. Mas é uma perspectiva, para além da dureza, uma perspectiva interessante. (Entrevista com dirigente de ADL – 26)

Já as **auditorias externas promovidas por entidades nacionais no âmbito de programas cofinanciados por fundos comunitários são consideradas menos exigentes** ao nível do acompanhamento da intervenção e sobretudo associadas às vertentes financeira e pedagógica (quando se trata de entidades formadoras). O enfoque é na “conformidade técnico-legal dos *dossiers*” e é dada pouca atenção à implementação e aos resultados das intervenções.

As auditorias são realizadas pontualmente (...). Podem vir diretamente do POPH ou do IGFSE e incidem sobre as duas vertentes: pedagógica e financeira. Vêm avaliar o próprio dossier técnico-pedagógico, a forma com está instruído, se incorpora toda a documentação indispensável e definida como a necessária para validar todos estes processos, e depois, do ponto de vista financeiro, se os documentos estão todos devidamente identificados e se são ajustados com base em critérios de razoabilidade. É um a avaliação administrativa e financeira (Entrevista com direção executiva de escola profissional – 15)

Os resultados do inquérito aproximam o perfil das organizações com processos de certificação ao perfil de organizações com processos de auditoria (especialmente de âmbito técnico), nomeadamente: a sua particular associação aos serviços sociais, educativos, de saúde ou de desenvolvimento, ao financiamento público, a organizações de grande dimensão financeira (sobretudo ADL e IPSS); e a sua dissociação de entidades de constituição recente. Esta proximidade de perfis tem que ver com a **sobreposição de lógicas dos dois tipos de processos**. Desde logo, o facto de a auditoria ser uma componente central nos processos de certificação e de estar muito associada à prestação de serviços sociais subsidiados pelo Estado (particularmente no âmbito de visitas de acompanhamento técnico no âmbito de acordos de cooperação). Por outro lado, estando os processos de

auditoria externa particularmente vocacionados para a verificação da legalidade e regularidade da aplicação dos recursos, é compreensível que as entidades com maior dimensão financeira e volume de financiamento público – como é o caso das IPSS e ADL da amostra – sejam as mais sujeitas a este tipo de ação de aferição por parte das entidades tutelares/financiadoras.

4.3.4 Compreendendo a implementação de práticas de gestão, controlo e avaliação: perfis de organizações do terceiro setor

A análise das **características organizacionais influentes das diversas práticas de gestão e avaliação** que vem sendo feita ao longo da análise, permitindo-nos esboçar o tipo de perfil organizacional que mais se associa a determinado tipo de prática avaliativa, deixa a descoberto não só a **forte influência dos atributos internos das organizações** (como seja a dimensão financeira, a antiguidade e o tipo de funções/papéis), **mas também das forças externas**, nomeadamente as associadas às relações de financiamento, na adoção de práticas avaliativas.

Os exercícios de análise (univariada) conduzidos até agora têm-nos permitido conhecer o efeito daqueles fatores por si só, individualmente considerados. Todavia, esta análise não nos fornece informações no sentido de uma perspetiva dinâmica de interinfluência, delineando-nos um panorama estático dos efeitos. A presença de vários fatores em simultâneo poderá, por um lado, aumentar a influência de um dado fator ou mesmo evidenciar uma importância não registada isoladamente, por outro anular ou reduzir a influência de fatores isoladamente significativos, mas não na presença de um conjunto de fatores de mais elevado poder explicativo⁸⁸. Dada a extensão da análise realizada, materializada na criação de **um modelo para cada uma das práticas (14**

⁸⁸ Para a construção de cada um dos modelos foi incluído o conjunto de variáveis anteriormente interpretadas na análise univariada, na medida em que foram encontradas influências significativas destas para a generalidade dos exercícios de gestão, controlo e avaliação listados. Procurou-se construir um modelo geral explicativo do recurso aos diferentes exercícios de gestão e avaliação, modelo este que, adequado aos dados, incorpore os fatores relativos à estrutura organizativa que, intercorrelacionados, sejam significativos para a prevalência do recurso a cada uma das práticas listadas. Optou-se pela utilização de modelos de regressão categorial ordinal. A escolha de função *Link* foi feita de acordo com os critérios de distribuição de frequências de classes da variável dependente definidos em Maroco (2014). O pressuposto do modelo da homogeneidade dos declives foi avaliado em todas as análises.

modelos), apresentam-se apenas as conclusões gerais da análise, sem explanação exaustiva dos resultados⁸⁹. Conclui-se que:

- As regressões reforçam os resultados ao nível dos testes utilizados para a comparação dos diversos grupos de entidades. Com efeito, a **dimensão** está no **topo das variáveis influentes** da maioria dos processos avaliativos, revelando a sua maior prevalência nas organizações com **maiores recursos**, mas apresentado um **poder preditivo relativamente mais modesto** nos processos de monitorização e avaliação associados ao projetos (**relatórios e avaliações formais**), nos quais a distância entre médias e grandes organizações é atenuada.
- Nos modelos estatísticos concebidos a variável **tipo de entidade** assume também grande **centralidade na generalidade dos processos, muito por conta da presença das coletividades**, que se distanciam significativamente das outras organizações na adesão a quase todas as práticas. A dupla **ONG/ADL** destaca-se nos processos de relatório e de **avaliação associados aos projetos** (modelos lógicos, relatórios financeiros/de atividades), enquanto a **dupla IPSS/ADL** sobressai ao nível da **certificação e da auditoria**.
- A **principal fonte de receitas** também está no topo das variáveis influentes da maioria das práticas, à exceção da monitorização de objetivos organizacionais e relatórios anuais. As regressões fazem sobressair a influência das receitas provenientes de **filantropia** e de **fundos públicos** (especialmente os de carácter temporário - **projetos**) na **maior adesão** a generalidade dos processos. Não obstante, as entidades **dependentes de financiamento público destacam-se** significativamente das primeiras na **adesão a processos de certificação, auditoria e relatórios de projetos**.

⁸⁹ À exceção do relativo à elaboração de relatórios anuais (contas/planos de atividades), todos os modelos criados são estatisticamente melhores que os modelos nulos, isto é, pelo menos um dos fatores inseridos nos modelos afeta significativamente as probabilidades de ocorrência de cada uma das práticas. Não obstante, o poder explicativo das regressões realizadas é genericamente baixo ou moderado; ou seja, os modelos criados permitem explicar percentagens modestas da variância dos processos de avaliação. Os modelos em que a dimensão dos efeitos é mais forte são os relativos aos processos de controlo da qualidade e auditoria e de avaliação de projetos – explicam entre cerca de 30 % a 40% da variância. Os resultados estatísticos mais modestos, por seu turno, são observados nos modelos criados para ferramentas específicas de avaliação (SROI, “balanced scorecard”, modelos lógicos) e para processos de monitorização e avaliação do desempenho à escala organizacional – explicam entre cerca de 12% a 19% da variância.

- Numa análise global das regressões realizadas para cada um dos processos de gestão e avaliação, o **tipo de funções**, a par com a **antiguidade**, são as variáveis com **menor influência** na predição das diversas iniciativas. A presença de vários fatores em simultâneo (de mais elevado poder explicativo) terá reduzido a influência destes fatores, isoladamente significativos. Ainda assim, é possível observar que o tipo de **funções/papéis** é particularmente **influyente da utilização de ferramentas específicas de medição e avaliação** (SROI, modelos lógicos e BSC), por via da dissociação de entidades funções exclusivamente de expressão. No que toca à antiguidade, sobressai a tendência para a **dissociação das organizações mais antigas** da maioria dos processos.

O quadro que se segue sintetiza os resultados das análises univariadas e multivariadas. Assim, associa-se a cada uma das práticas o perfil de organização - em termos das suas características – em que se observa uma maior adesão, apresentando-se apenas aquelas para as quais se verificam diferenças estatisticamente significativas (*p-value* > 0,005). As características destacadas a negrito são as evidenciadas na análise multivariada, correspondendo às variáveis mais influentes nos modelos de regressão categorial.

Quadro 9. Práticas de gestão, controlo e avaliação: perfis de OTS

	DIMENSÃO	TIPO	RECEITAS	FUNÇÕES	IDADE	TOTAL*
	ENTIDADE					
Reporte						
Relatórios anuais	Grandes/ Médias	-	Públicas/ Próprias	Serviços	-	1
Relatórios atividades de projetos	Grandes/ Médias	ONG/ADL	Públicas	Serviços/ Mistas	Intermédia/ Recente	4
Relatórios execução financeira de projetos	Grandes/ Médias	ONG/ADL	Públicas	Serviços/ Mistas	Intermédia/ Recente	3
Controlo de qualidade e auditoria externa						
Certificação	Grandes	ADL/IPSS	Públicas (projetos)	Serviços	Intermédia	4
Auditorias financeiras externas	Grandes	ADL → IPSS (< Coletiv.)	Públicas (projetos)	Serviços	Intermédia/ Recente	3
Auditorias técnicas externas	Grandes	ADL/IPSS	Públicas	Serviços	Intermédia	3

Avaliação						
Avaliações formais de projetos	Grandes/ Médias	ADL/ONG (< Coletiv.)	Filantropia/ Públicas (projetos)	Serviços/ Mistos	Recente/ Intermédia	4
Análises SROI – Social Return on Investment	Grandes	ADL/IPSS/ ONG (< <i>Outras</i>)	(< <i>Próprias</i>)	Serviços/ Mistos	-	4
Conceção de “modelos lógicos”	Grandes/ Médias	ONG/ADL	(< <i>Próprias</i>)	Serviços/ Mistos	Recente/ Intermédia	5
Avaliações formais desemp. organizac.	Grandes	(< Coletiv.)	Filantropia/ Públicas	Serviços/ Mistos	-	3
Gestão estratégica e do desempenho						
Avaliação desemp. pessoal ao serviço	Grandes	IPSS/Outras (< <i>Coletiv.</i>)	Filantropia/ Públicas	Serviços	Recente/ Intermédia	3
Monitoriz.objetivos e metas organizac.	Grandes	(< <i>Coletiv.</i>)	Filantropia/ Públicas	Serviços/ Mistos	Recente/ Intermédia	3
Elaboração de plano estratégico	Grandes	(< <i>Coletiv.</i>)	Filantropia/ Públicas (projetos)	Serviços/ Mistos	Intermédia/ Recente	4
<i>Balanced scorecard</i>	Grandes	IPSS/ONG/ ADL	Filantropia/ Públicas (projetos)	Serviços/ Mistos	Intermédia/ Recente	4
Total**	12	12	12	5	7	

Fonte: Inquérito por questionário.

Notas: * N.º total de variáveis independentes (dimensão, tipo de entidades, ...) que revelam poder preditivo na análise multivariada.

** N.º total de casos em que a variável independente revela poder preditivo do comportamento da variável dependente na análise multivariada.

Não obstante a relevância de atributos geralmente conotados com a dimensão interna das organizações, como é o caso da dimensão financeira, a análise é reveladora da **influência das pressões institucionais associadas à dependência de recursos** (aqui sobretudo indicadas pelo perfil de financiamento) **na configuração das práticas das OTS**, coerentemente com os pressupostos da teoria da agência e da dependência de recursos⁹⁰.

⁹⁰ A relevância das fontes de recursos na compreensão da adesão à avaliação é reforçada pelos resultados do modelo explicativo da “**intensidade/abrangência das atividades de avaliação**” (variável medida numa escala ordinal de 0 – Não fazemos qualquer avaliação das nossas atividades; a 4 – Monitorizamos e avaliamos todas as nossas atividades). Recorreu-se à regressão ordinal que, para além da principal fonte recursos, permitiu integrar outras variáveis que descrevem características da estrutura organizacional influentes das práticas de gestão e avaliação (dimensão, tipo de entidades, funções principais e antiguidade), ponderando os respetivos efeitos no modelo. O modelo é estatisticamente significativo ($p=0,003$), ainda que a dimensão dos efeitos seja reduzida ($R^2_N = 0,060$). No modelo, o **tipo de receitas** está no **topo da influência** da variável dependente (intensidade da avaliação), por conta da maior probabilidade

O vigoroso efeito das forças externas na decisão de avaliar é, desde logo, evidenciado pelas tendências verificadas nas organizações que serão mais imunes a pressões coercivas das fontes de recursos – as organizações autossustentadas. Vemos estas entidades se destacam pela menor adesão a generalidade das atividades listadas. Por seu turno, as entidades cujas receitas advêm sobretudo da **filantropia individual ou coletiva** são as que mais se vinculam aos exercícios de avaliação mais compreensivos e estruturados, quer à escala dos projetos quer à escala organizacional (desvinculando-se de exercícios de natureza mais normativa/coerciva, como o controlo e auditoria externos e o reporte descritivo ao nível dos projetos)⁹¹. A este tipo de de práticas conotadas com a prestação de contas ascendente e com as pressões coercivas exercidas pelas fontes de recursos encontram-se particularmente vinculadas as entidades dependentes do financiamento público, especialmente o de carácter temporário (através de **projetos nacionais ou comunitários**). Mas, enquanto as entidades sobretudo sustentadas por **subsídios** apenas sobressaem na adesão a este tipo de práticas, as entidades com financiamento público por projeto destacam-se pela associação a uma maior diversidade de processos, nomeadamente a avaliação formal de projetos⁹², parecendo revelar, grosso modo, uma **maior cultura de avaliação**. No entanto, é também perceptível a debilidade da influência deste modelo de financiamento nas práticas estruturadas de avaliação e medição do desempenho e do impacto **de nível organizacional** (e individual). Estas constatações apontam para a ideia de que a avaliação conduzida em organizações

das classes de menor ordem (**menor intensidade**) nas organizações **autossustentadas** ($p=0,028$) com características semelhantes em termos de dimensão, antiguidade, natureza e perfil de atividade. Além das fontes de receita, apenas o **tipo de entidade** exerce influência no modelo, por conta da maior probabilidade das classes de maior ordem (maior intensidade) nas **ONG** ($p=0,030$).

⁹¹ As práticas de **reporte descritivo** sobre a execução física e financeira dos **projetos**, são as únicas práticas de gestão e avaliação elencadas que se associam (positivamente) e de forma estatisticamente significativa à **adoção coerciva das práticas de avaliação**, o que é relevado pela particular adesão à opção “Fazemos avaliação porque os financiadores exigem” no exercício estatístico que correlaciona as diversas práticas de gestão com um conjunto de variáveis relativas a perceções e práticas de avaliação associadas às fontes de financiamento. Já as **práticas de monitorização e avaliação mais sistemáticos e formais**, sobretudo à escala organizacional, tendem a dissociar-se de forma estatisticamente significativa da adesão à avaliação por imposição (vide correlações na tabela 1 do anexo 5).

⁹² As práticas avaliativas de natureza compreensiva e analítica/crítica à escala dos projetos (avaliações formais, SROI e ML) tendem a escapar à visão da avaliação enquanto fardo imposto pelas entidades financiadoras e a sua associação à sobrevivência organizacional, dissociando-se de afirmações como “fazemos avaliação porque os financiadores exigem” ou “fazemos avaliação para obter novos financiamentos” (vide correlações na tabela 1 do anexo 5).

dependentes do financiamento de projetos poderá ser particularmente **vulnerável a imperativos de curto prazo de prestação de contas externa** e a necessidades de informação das entidades financiadoras, sem verter para os processos mais amplos de análise do desempenho organizacional, ancorados no cumprimento da missão⁹³. Tal ideia reenvia aos **constrangimentos do modelo de financiamento por projeto** nas organizações do terceiro setor assinalados noutras análises⁹⁴, nomeadamente a sobreposição dos interesses das entidades financiadoras às necessidades das organizações candidatas, e o facto de o ciclo dos projetos ser muitas vezes curto e desalinhado da estratégia geral das organizações. Esta tendência é reforçada pelos resultados estatísticos relativos à **influência da participação em projetos** (sem que estes constituam necessariamente a maior base de receitas das entidades) na adesão às práticas de gestão e avaliação⁹⁵. Apesar de a participação em projetos ser preditora do reforço do envolvimento num conjunto mais alargado de processos conotados com a análise do desempenho, quando comparada com o efeito do modelo de financiamento por projeto, os únicos exercícios em que a sua influência não tem relevância estatística correspondem exatamente àqueles em que a influência do modelo de financiamento por projeto é menos significativa (i.e., nos processos de monitorização e avaliação formal à escala organizacional).

Por outro lado, podemos observar que a participação em projetos comunitários⁹⁶ tem um **efeito positivo na mitigação da distância** entre organizações determinada pela

⁹³ Análises anteriores (por exemplo, Ebrahim, 2005; Ellis e Gregory, 2008; Hoole e Patterson, 2008) já alertavam, exatamente, para o perigo de a prestação de contas a financiadores/as, operacionalizada através de relatórios ou outras formas de controlo, poder estar em desalinho com a responsabilização perante a missão das organizações.

⁹⁴ Estes constrangimentos, aludidos em análises nacionais e internacionais, encontram-se descritos no primeiro ponto do segundo capítulo.

⁹⁵ Para avaliar se o envolvimento em projetos apresenta um efeito estatisticamente significativo sobre as probabilidades de adesão a cada uma das práticas de controlo e avaliação listadas, recorreu-se à regressão ordinal (14 modelos), ponderando também o efeito de outras variáveis influentes na adesão (tipo de organização, funções principais, dimensão e antiguidade). A participação em projetos revela influência estatisticamente significativa em 10 dos 14 modelos concebidos, com *p-values* a variar entre 0,000 e 0,041.

⁹⁶ Apesar de no inquérito se diferenciar a participação em projetos POPH e EQUAL, **a generalidade das OTS inquiridas que indicam a participação em projetos EQUAL estão ou já estiveram também envolvidas em projetos POPH (43 em 51 entidades), pelo que o isolamento do efeito da participação naquele programa na análise dos processos de avaliação explorados no inquérito não se revela profícuo** em termos de significância. Por essa razão, e embora a comparação entre os dois programas constitua um eixo relevante da análise, ela não poderá ser feita com rigor a partir deste tipo de desagregação, pelo que os resultados serão lidos globalmente.

dimensão financeira, eventualmente por ela possibilitar a afetação de recursos ao desenvolvimento de exercícios de monitorização e avaliação (pelo menos os mais básicos) sem que as entidades tenham que recorrer a fundos próprios.

Além do perfil de financiamento das OTS, também o **subsetor de atividade compõe o ambiente externo em que as organizações operam**. Diferentes domínios de atividade estão associados, por exemplo, a ambientes de financiamento diversos (diferentes fluxos e requisitos) – de que decorrem pressões **coercivas** distintas –, a diferentes níveis de institucionalização e autorregulação e a distintas redes de organizações – de que resultam pressões **normativas** e **miméticas** diversas. São, aliás, vários os estudos que têm vindo a salientar diferenças entre subsectores/campos/domínios de atividade nos padrões de adoção (e uso) da avaliação e medição do desempenho/impacto (Hall *et al.*, 2003; Carman, 2007; Ellis and Gregory, 2008; Salamon, Geller e Mengel, 2010; Lumley *et al.*, 2011).

A análise da influência do âmbito de atividade sobre a frequência do envolvimento nas diferentes atividades de gestão, controlo e avaliação, sistematizada em anexo (vide tabela 2, anexo 5), é reveladora de algumas *nuances* no tipo de práticas adotadas nos diferentes subsectores do TS, mesmo comparando organizações com dimensões similares (através da controlo da variável “dimensão”⁹⁷ na análise multivariada):

- A **monitoria externa** está particularmente vinculada aos setores da educação, desenvolvimento e ambiente, serviços sociais, participação cívica e defesa de causas e associações empresariais, profissionais e sindicais.
- As práticas de **avaliação estruturada à escala dos projetos** são sobretudo associadas aos setores da cooperação/desenvolvimento internacional, do desenvolvimento local e ambiente, e da participação cívica e defesa de causas.
- O **reporte descritivo** ao nível dos projetos é típico de setor do desenvolvimento local e ambiente.

⁹⁷ A dimensão das organizações distingue muito os subsectores de atividade: as organizações no setor da cultura/lazer e participação cívica e defesa de causas destacam-se pela sua menor dimensão (88% e 69%, respetivamente, têm pequena ou muito pequena dimensão).

- A **monitorização** de objetivos e metas organizacionais associa-se particularmente ao setor da atividade internacional.

Dentre uma diversidade de lógicas institucionais a que se podem associar processos de adaptação/isomorfismo o campo organizacional, o estudo ressalta também os efeitos da **pertença a redes de organizações**. O contacto com outras organizações e profissionais, potenciado, designadamente, pela pertença a redes de organizações (em particular, à **escala internacional**) é evidenciado quer em contexto de entrevista, quer nos resultados do inquérito, assumindo um lugar importante no estímulo à adoção de práticas de monitorização e avaliação. Com efeito, nas **organizações que afirmam participar em redes/organismos internacionais não só a intensidade das atividades de avaliação é tendencialmente maior**, em termos globais⁹⁸, como ela é superior na generalidade das práticas, analisadas individualmente (vide tabela 3, anexo 5). Essa influência é particularmente marcada nas práticas de monitorização e avaliação enquadradas no planeamento estratégico e no de reporte e avaliação decorrentes do envolvimento em **projetos**, em especial a conceção de modelos lógicos (muito utilizada na cooperação para o desenvolvimento).

Se os resultados do estudo dão suporte à relevância das forças miméticas e normativas decorrentes da pertença a organismos/redes internacionais, o efeito da pertença a **organizações de cúpula** no impulso das práticas de monitorização e avaliação nas organizações – sendo potencialmente relevante pelo papel normativo sobre as associadas e de autorregulador do sector – não é tão eloquente. Com efeito, a pertença a plataformas nacionais representativas não exerce um efeito estatisticamente significativo nem na intensidade/abrangência da avaliação nem na adesão a qualquer das práticas elencadas.

⁹⁸ Para avaliar se a intensidade da avaliação (variável medida numa escala ordinal de 0 – Não fazemos qualquer avaliação das nossas atividades; a 4 – Monitorizamos e avaliamos todas as nossas atividades) varia de acordo com a pertença a redes/organismos internacionais recorreu-se ao teste não paramétrico de U de Mann-Whitney. Usou-se uma probabilidade de erro de tipo I (α) de 0,05. Vemos que a pertença a organismos internacionais exerce um efeito estatisticamente significativo sobre o grau de intensidade das práticas de monitorização e avaliação (*p-value*= 0,02).

4.4. Decompondo as práticas avaliativas: métodos, conteúdos e intervenientes

4.4.1. Metodologias de avaliação: entre a informalidade das abordagens qualitativas e apreciativas e a conformidade com os requisitos de descrição e quantificação

Nos debates sobre abordagens/metodologias mais adequadas para avaliar a intervenção das organizações colocam-se geralmente em confronto dois registos diferentes de desenvolvimento metodológico decorrentes da natureza da avaliação enquanto investigação aplicada: o quantitativo e o qualitativo (vide quadro 10). Estas discussões, que frequentemente se centram na superioridade de umas técnicas sobre outras, contrapõem os seguintes tipos de argumentos: a pesquisa qualitativa “não é científica, é subjetiva, limitada e baseada em técnicas brandas”, e a pesquisa quantitativa “é superficial, alheia da realidade, submetida a um empirismo estéril e baseada em técnicas duras”. Os conflitos sobre a utilização de diferentes tipos de dados (quantitativos *versus* qualitativos) são, pelo menos em certa medida, manifestações de desacordos sobre o que constitui um processo de avaliação ideal (Hall, 2014).

Quadro 10. Características dos paradigmas quantitativo e qualitativo

Paradigma quantitativo	Paradigma qualitativo
<ul style="list-style-type: none">▪ Uso de métodos quantitativos	<ul style="list-style-type: none">▪ Uso de métodos qualitativos
<ul style="list-style-type: none">▪ <i>Positivismo</i> lógico	<ul style="list-style-type: none">▪ <i>Fenomenologia</i> e <i>compreensão</i>, preocupação com a perspetiva do ator
<ul style="list-style-type: none">▪ Mensuração obstrutiva e controlada	<ul style="list-style-type: none">▪ Observação naturalista e controlada
<ul style="list-style-type: none">▪ Objetividade	<ul style="list-style-type: none">▪ Subjetividade
<ul style="list-style-type: none">▪ Perspetiva desde fora, distante dos dados	<ul style="list-style-type: none">▪ Perspetiva interna, próxima dos dados
<ul style="list-style-type: none">▪ Orientado para a verificação e confirmação das hipóteses, <i>hipotético-dedutivo</i>	<ul style="list-style-type: none">▪ Orientado para a descoberta de hipóteses, exploratório, indutivo, descritivo
<ul style="list-style-type: none">▪ Orientado para resultados	<ul style="list-style-type: none">▪ Orientado para processos
<ul style="list-style-type: none">▪ Dados confiáveis e replicáveis	<ul style="list-style-type: none">▪ Dados válidos, “ricos” e “profundos”
<ul style="list-style-type: none">▪ Particularista	<ul style="list-style-type: none">▪ Holístico
<ul style="list-style-type: none">▪ Assume uma realidade estável	<ul style="list-style-type: none">▪ Assume uma realidade dinâmica

Fonte: Cano, 2006: 90.

Um desentendimento comum diz respeito ao argumento de que os estudos de caso e histórias de vida podem ser subjetivos, enquanto os indicadores de desempenho e as estatísticas são mais objetivos e fiáveis. Nos antípodas, continuam a ser levantadas

questões sobre se as medidas quantitativas, em particular as financeiras, são por si só adequadas para captar e medir o desempenho e, sobretudo, o impacto social do terceiro setor. Enquanto o valor económico de uma qualquer atividade pode ser facilmente medido através de indicadores financeiros, o valor social de uma intervenção inclui benefícios intangíveis que desafiam a sua mensuração.

Esta última perspetiva é prevalecte entre as pessoas das OTS entrevistadas, sobretudo entre pessoas de entidades classificadas nos estudo como ONG, para quem os indicadores quantitativos não permitem captar aspetos qualitativos essenciais da intervenção social, ou que, devido à sua ênfase na quantificação, podem deturpar ou subestimar os aspetos qualitativos que lhe estão subjacentes. Depois, **a variedade do impacto social criado pelo terceiro setor torna difícil captar todos os tipos de impactos de forma justa e objetiva**. Os excertos que se seguem são reveladores deste ponto de vista:

Isto da avaliação também se nota numa questão que não sei se é muito positiva: a necessidade de cada vez maior quantificação dos resultados. Uma maior quantificação de resultados permite uma avaliação de resultados menos aberta, muito mais fechada. Mas, por outro lado, perde-se uma quantidade de informação e de questões que não são quantificáveis! (...) Se, por um lado, nos permite ganhar uma maior objetividade quando analisamos os resultados, por outro lado, faz-nos perder claramente o melhor ou o pior da implementação de um projeto. (...) (Entrevista com dirigente de ONG – 6)

Há atividades com poucas frequências mas com grande impacto. (...) Há por vezes absurdos na quantificação dos resultados. Em todo o resto não há nenhuma avaliação que eu acho mal. (Entrevista com técnico/dirigente de ONG – 31)

Atores internacionais relevantes no campo da avaliação, como sejam a UE, o Banco Mundial, a OCDE, a Associação Americana de Avaliação e a Sociedade Europeia de Avaliação (Furubo *et al.*, 2002; Stern, 2006), exaltam as mais-valias das avaliações multi-métodos (que combinam diferentes instrumentos e técnicas), **parecendo existir um amplo consenso de que as metodologias quantitativas e qualitativas podem ser combinadas com sucesso**. Parece prevalecer a ideia de que se deve articular e combinar avaliações que resultem em abordagens mais pluralistas e contextualizadas, que procurem fazer uma leitura mais ampla da realidade. Nas abordagens quantitativas, predomina uma maior preocupação com o sistema de ação movido pelas intervenções, no sentido da quantificação de metas e resultados, procurando mensurar o social por meio de índices numéricos. Já nas abordagens qualitativas enfatiza-se a lógica estabelecida pelos diferentes atores que movem a intervenção (coordenadores, executores/as, parceiros, financiadores,

peças beneficiárias, etc.) que imprimem direção e dinâmica à ação. Nas abordagens pluralistas ou ecléticas, por seu turno, procura-se uma inter-relação entre métodos qualitativos e quantitativos, isto é, não medem apenas quantitativamente os benefícios ou malefícios de uma intervenção, mas também qualificam as decisões, processos, resultados e impactos (Carvalho, 2000).

Nos guias e orientações para a avaliação de alguns destes organismos é comum encontrar vias para a utilização de abordagens diversas, **sem prescrição de uma metodologia de avaliação padrão**. Por exemplo, no guia para a avaliação do desenvolvimento económico (ELVASED), que informa a avaliação no âmbito dos Fundos Estruturais Europeus, pode ler-se:

Todos os métodos e técnicas têm pontos fortes e fracos (...). Para qualquer avaliação, as técnicas deverão ser escolhidas e aplicadas de forma a explorar as suas virtudes e a reconhecer as suas limitações. Face ao exposto, o melhor é aplicar métodos e técnicas em combinação como parte de qualquer trabalho de avaliação específico. O apoio num só método ou técnica será mais frágil do que a obtenção de múltiplas perspetivas (a por vezes denominada triangulação). (CE, 2013: 74)

No guia para a avaliação de impacto do Banco Mundial também encontramos uma clara adesão às metodologias ecléticas, como sobressai no seguinte excerto:

Recorra a um plano de avaliação multi-método. Tenha em mente as vantagens comparativas dos diferentes desenhos e métodos. A associação de métodos pode ser empregue para avaliar diferentes facetas de impactos ou resultados complexos, garantindo maior profundidade, largura e extensão no retrato criado, do que poderá garantir um único método. (...) (BM, 2010: 39)

Mas ao mesmo tempo que se reconhece a pertinência das abordagens qualitativas, **parecem privilegiar-se as lógicas da medição e da quantificação, sobrepondo-se a importância da quantificação das realizações e dos resultados**, mediante o acompanhamento de indicadores de desempenho, **no enquadramento da gestão por objetivos e orientada para os resultados**. O objetivo que parece prevalecer é o de recolher informação que seja comparável e permita, de forma objetiva e inequívoca, determinar se os objetivos estabelecidos e resultados esperados das intervenções são alcançados. No entender das entidades promotoras no terreno, em geral, as entidades financiadoras (nacionais e internacionais), privilegiam a quantificação dos resultados. A informação qualitativa é frequentemente entendida como “suplemento” das análises quantitativas.

É preciso que alguma coisa se altere e que as metodologias qualitativas voltem a aparecer com uma força muito grande na avaliação de projetos e da avaliação

das organizações. Enquanto os indicadores estiverem todos preparados para ser quantitativos, porque vivemos num mundo de números, vai ser uma avaliação feita até remotamente em função da capacidade que as organizações têm de provar que fizeram as atividades. (...) Obviamente que tem que ser quantificado, não tem é que ser só quantificado. (Entrevista com dirigente de OTS – 6)

Associadas à distinção entre abordagens qualitativas e quantitativas estão os enfoques **descritivo** e **apreciativo/crítico** da avaliação. Como desde há muito defende Robert Stake, a avaliação contém sempre estas duas vertentes – “Penso no ato de julgamento como parte do ato descritivo, como parte do ato de observação” (Stake, 2001: 10). Observar e julgar são, nesta ótica, parte do mesmo ato, sendo a tarefa da avaliação tanto uma questão de refinamento das primeiras perceções sobre a qualidade de uma intervenção como de construção de um corpo de evidências para determinar o nível de qualidade (Christie e Alkin, 2013: 34). Na linha desta perspetiva, as avaliações sólidas e úteis, como as designam Stufflebeam e Coryn (2014: 16), baseiam-se em informação **descritiva e crítica**, na medida em que descrevem os processos e os resultados inerentes às intervenções, ao mesmo tempo que fazem **apreciações** sobre o seu valor ou mérito. A informação descritiva, traduzindo factos, atribui substância à informação apreciativa, a qual poderá incluir o *feedback* orientado para a valorização e para a identificação de pontos fortes e fracos das intervenções.

Apesar de, por princípio, nas avaliações coexistirem as vertentes descritiva e crítica, é frequente elas serem acopladas a tipos de avaliação distintos, consoante a centralidade de um ou de outro tipo de informação. Nesta linha, **as avaliações descritivas são associadas às medições quantitativas**, à obtenção de parâmetros quantitativos que podem ser aplicados a um grupo de indivíduos que beneficiam ou participam num programa ou intervenção (Cohen e Franco, 2006: 116). Uma vez que as avaliações da implementação dos programas ou projetos procuram informação sobre a população coberta pelas intervenções e sobre as respetivas atividades, tendem a ser geralmente de natureza descritiva (Cuesta e Martínez, 2010). A **avaliação analítica/crítica é associada à tentativa de responder às questões de causa e efeito**, requerendo o recurso a modelos teóricos e estratégias de investigação apropriados no sentido de demonstrar que interpretações causais são verdadeiras (Cohen e Franco, 2006: 116).

As distinções entre avaliação qualitativa/quantitativa e crítica/descritiva aparecem com relevância no estudo. A **fundamentação das avaliações em informação quantitativa e descritiva** sobre outputs e resultados parece ser particularmente marcada (e contestada) nas **avaliações sumativas de projetos realizadas por iniciativa das entidades financiadoras**, como é o caso das auditorias realizadas por agências financiadoras internacionais no âmbito de projetos de cooperação para o desenvolvimento. Parece também ser o caso das conduzidas pela equipa do programa “Escolhas”, cujos critérios de classificação dos projetos são fortemente contestados pelas instituições promotoras entrevistadas, como no caso que se traz à colação:

Os conceitos de avaliação, a partir do momento em que se quer exagerar na quantificação... Quantos telefonemas foram feitos, quantos km foram andados... isto é um absurdo. Isto altera tudo o que são as dinâmicas do próprio trabalho. (...) Chegou-se a um momento em o projeto começou a fazer coisas para os números se não para as pessoas. E aí finalmente teve um bom. (...) Há quantificações que não levam à avaliação, (...) antes pelo contrário, desvirtuam avaliação. (...) Há atividades com poucas frequências mas com grande impacto. (Entrevista com dirigente/técnico de OTS – 31)

No caso dos programas selecionados para estudo mais aprofundado - EQUAL e POPH - o tipo de abordagens suscitadas pelos requisitos de avaliação ao nível dos projetos tendem a enquadrar a avaliação conduzida pelas organizações promotoras em **modelos distintos**.

1. No caso do **PIC EQUAL**, em termos metodológicos, quer ao nível da autoavaliação quer ao nível da avaliação independente, **estimulava-se a combinação de metodologias e técnicas de avaliação “quantitativas” e “qualitativas”, e de informação descritiva e apreciativa**, que dessem conta da singularidade dos processos e dos resultados dos projetos. A importância da conjugação de metodologias qualitativas e quantitativas e de informação descritiva e analítica foi visibilizada tanto nos guias e nas sessões de apoio à avaliação nos projetos como nas orientações para a avaliação independente dos projetos EQUAL. Embora fosse requerida a recolha de dados quantitativos sobre a execução física e financeira dos projetos, fomentava-se a sua articulação com a análise intensiva de informação qualitativa, priorizando-se a análise aprofundada em detrimento de uma avaliação extensiva e partilhada pelos diversos atores envolvidos nas diferentes dimensões dos projetos (Figueirôa e Pimenta, 2005). Por outro lado, ao mesmo tempo que estimulava a recolha de informação descritiva sobre a implementação (por exemplo, n.º de atividades

previstas e realizadas e n.º de destinatários/as previstos e verificados), também promovia a **análise crítica e sistemática dos dados obtidos** (por exemplo, através da colocação de questões às parcerias como "avalie o progresso das atividades/n.º de destinatários face ao esperado"), premindo "que as pessoas e as instituições envolvidas [julgassem] o seu trabalho e [aprendessem] com os julgamentos feitos" (Gabinete de Gestão EQUAL, 2003: 11). A importância das **apreciações sobre o mérito e valor das intervenções** era particularmente afirmada nas orientações do gabinete de gestão do PIC EQUAL para a avaliação independente dos projetos, prescrevendo para o exercício de avaliação final a adoção de uma perspetiva de "avaliação sumativa que determine até que ponto os objetivos foram efetivamente alcançados e produza um juízo de valor sobre os resultados obtidos" (Gabinete de Gestão EQUAL, 2005: 3).

2. No caso do POPH era apenas obrigatória a recolha de **dados quantitativos** de carácter **descritivo/factual** para alimentar os indicadores de realização e resultado considerados relevante para aferir sobre o **desempenho do próprio programa na concretização dos seus objetivos**. Tratava-se sobretudo de dados administrativos (número de atividades, tipos de intervenções realizadas, horas de formação, número de pessoas que concluíram ações, n.º de postos de trabalho criados) recolhidos pelas organizações de forma rotineira e que não convocam a utilização de técnicas qualitativas, como entrevistas, *focus group* ou estudos de caso. O enfoque do programa era nos "números":

Na plataforma faz-se a avaliação, mas eles olham apenas para no nº de destinatários, volume de formação, para aquelas dados físicos que temos que colocar na plataforma, e se aquilo está de acordo com análise financeira, fazem a comparação da análise financeira com a física no sentido de verificar se o valor que foi gasto está de acordo com os dados físicos, não importa (...). Ter 12 formandos, mas entretanto ter-se reduzido para 8, mas se calhar achou-se melhor porque ia dar-se melhor resposta e que em termos de impacto seria mais efetivo do que se tivéssemos os 12. Para eles, isto não interessa. (Entrevista com dirigente de ONG – 27)

Esta **informação descritiva, traduzindo factos, poderia servir de base à formulação de juízos** sobre o desempenho dos projetos nos planos da operacionalização, realização e efeitos imediatos. Todavia, como não era obrigatória a apresentação de relatório de avaliação, a contextualização e atribuição de significado e de valor aos dados quantitativos e a compreensão dos processos associados os outputs e resultados conseguidos teve um alcance muito limitado.

Análises estatísticas das relações entre a participação em projetos e o tipo de informação de monitorização e avaliação mais frequentemente recolhida nas organizações inquiridas reforçam a ideia da **prevalência do paradigma quantitativo na avaliação ao nível dos projetos**, assim como em ações de monitorização/acompanhamento da **iniciativa das entidades financiadoras/tutelares**. Com efeito, nas organizações envolvidas em projetos – em particular do POPH – e/ou em que o financiamento no âmbito de projetos é prevalectante na estrutura de receitas, a tendência para a primazia da informação quantitativa é reforçada⁹⁹. Por outro lado, ao estabelecer associações entre as diversas práticas de gestão, controlo e avaliação listadas com o tipo de informação privilegiada nas organizações inquiridas (vide tabela 5), vemos que as **práticas associadas aos projetos** (concretamente, relatórios de execução física e financeira e avaliações formais) – a par com as auditorias financeiras ou técnicas da iniciativa das entidades financiadoras – **são as que mais se associam estatisticamente à recolha de informação quantitativa**. Isto indicia que a **recolha de informação quantitativa apresenta um carácter mais coercivo** comparativamente à recolha de informação de natureza qualitativa, estando particularmente associada à **prestação de contas ascendente**¹⁰⁰.

⁹⁹ Para avaliar se o tipo de informação de avaliação recolhida é influenciada pela participação da organização em projetos recorreu-se ao Teste do Qui-quadrado de Pearson, considerando-se uma probabilidade de erro de tipo I de 0,05. A análise estatística permite afirmar que o tipo de informação de avaliação varia significativamente consoante a existência de projetos, em particular de projetos POPH, apontando para a prevalência da informação quantitativa (*p-value* = 0,000). Existem também diferenças estatisticamente significativas entre os tipos de informação priorizada (em favor da informação quantitativa) em organizações em que o financiamento no âmbito de projetos é dominante (*p-value* = 0,000). A participação em projetos EQUAL não exerce um efeito estatisticamente significativo no tipo de informação privilegiado (*p-value* = 0,225).

¹⁰⁰ Esta observação é reforçada pelo exercício estatístico de associação entre a priorização de informação quantitativa e um conjunto de variáveis relativas a perceções e práticas de avaliação associadas ao financiamento, verificando-se correlações positivas com relevância estatística entre o **predomínio da informação quantitativa e a condução de exercícios e avaliação por imposição das entidades financiadoras e para prestação de contas ascendente** (“Fazemos avaliação porque os financiadores exigem”/“Para reportar aos financiadores”). Tabela de correlações em anexo (vide tabela 4, anexo 5).

Tabela 5. Frequência do envolvimento em atividades de gestão e avaliação segundo tipo de informação de avaliação recolhida com maior frequência

Envolvimento em atividades de gestão e avaliação (1 – Nunca, 2 – Ocasionalmente, 3 – Regularmente)	Tipo de informação		Teste U de Mann-Whitney
	Informação quantitativa Média	Informação qualitativa Média	Significância Assint. (Bilateral)
Relatório anuais (contas/plano atividades)	2.9	2.9	,950
Relatórios atividades de projetos	2.6	2.4	,027
Relatórios execução financeira de projetos	2.5	2.3	,009
Realização de avaliações formais de projetos	2.3	2.2	,034
Processos de certificação/acreditação	1.8	1.7	,176
Auditorias financeiras externas	2.0	1.7	,001
Auditorias técnicas/visitas de acompanhamento	2.1	1.9	,005
Avaliação de desempenho do pessoal ao serviço	2.1	2.0	,583
Monitoriz. objetivos e metas organizacionais	2.5	2.4	,943
Elaboração de plano estratégico a 3 ou mais anos	2.0	2.0	,625
Sistema de gestão “balanced scorecard”	1.4	1.4	,053
Análises SROI – Social Return on Investment	1.4	1.5	,679
Avaliações formais desemp. organizacional	1.9	1.9	,385
Conceção de “modelos lógicos”	1.8	1.8	,816

Fonte: Inquérito por questionário.

A nível organizacional, as lógicas qualitativa e quantitativa parecem coexistir.

Como vimos atrás, a generalidade das organizações avaliam as suas atividades por intermédio de relatórios anuais de atividade, tendo por referência o planeamento anual da atividade e/ou, menos frequentemente, o planeamento estratégico. Em traços gerais, relatórios caracterizam-se, por uma lado, **pela presença conjunta de indicadores quantitativos e descritivos sobre a atividade da organização** (e.g., número de pessoas acompanhadas, número de atendimentos, número de formandos/as, número de folhetos distribuídos), alguns deles solicitados pela tutela ou por outras entidades financiadoras, e, por outro, por um conjunto de **informação qualitativa que contextualiza ou justifica os outputs e/ou descreve mudanças de mais longo prazo, e que assenta frequentemente em impressões do pessoal técnico e/ou fontes informais.**

A análise dos relatórios e planos anuais de atividade das OTS entrevistadas revela uma **forte tendência da generalidade das organizações em quantificar a sua intervenção, mas sobretudo ao nível das realizações/produtos/outputs das intervenções** (sobretudo

associados ao número de atividades/ações desenvolvidas e ao número de pessoas abrangidas) e muito menos ao nível dos efeitos/mudanças suscitados pela ação das organizações sobre as pessoas beneficiárias ou sobre a população mais vasta. Sendo os elementos mais tangíveis e objetivos do alcance da ação das organizações os *outputs* das intervenções, em geral descritos ou medidos com base em unidades físicas/materiais ou financeiras/monetárias, são mais facilmente mensuráveis e quantificáveis do que os efeitos de médio ou longo prazo. **O carácter imaterial e subjetivo de muitas das mudanças suscitadas pelas intervenções sociais** (por exemplo, a melhoria na autoestima, da autonomia ou do bem-estar) **desafia a sua quantificação**. Por outro lado, a mensuração de alguns efeitos **“objetivos” mais facilmente quantificáveis** (por exemplo, integração no mercado de trabalho após uma formação/estágio, diminuição do absentismo escolar, inserção profissional de jovens em risco) requer um acompanhamento e envolvimento dos públicos-alvo que frequentemente ultrapassa o período em que decorrem as atividades. Esta medição implicaria a mobilização de recursos para auscultar as populações decorrido determinado período após o fim da intervenção que muitas organizações não estão disponíveis para fazer. Ainda assim, encontramos exemplos (6 casos) de investimento em relação ao acompanhamento de alguns indicadores de resultado quantitativos que vão para além do número de ações realizadas ou do número de participantes nas ações, remetendo para mudanças de curto prazo nas competências, atitudes ou comportamentos. Apresenta-se de seguida o excerto do relatório de atividade de uma das organizações abrangidas pelo estudo em que a preocupação com o acompanhamento de alguns indicadores de resultado quantitativos é evidenciada:

Mais de 50% dos indivíduos acompanhados no programa de substituição opiácea, de forma sistemática, melhoram/mantiveram o acompanhamento médico e de saúde: 89 % têm análises atualizadas; 100% estão inscritos em Centro de Saúde; 100% dos indivíduos têm disponível acompanhamento médico e de saúde pelo médico e equipa de enfermagem a colaborar na unidade móvel; 25% são também acompanhados em consultas de especialidade na Unidade Local de Saúde de X; 100% dos utentes têm a vacinação monitorizada pela equipa de enfermagem. (Relatório de atividades de OTS)

A grande maioria dos relatórios de atividades caracteriza-se por um forte enfoque em informação descritiva, frequentemente com grande detalhe sobre os procedimentos adotados, mas sem apresentar visões sintéticas e de conjunto sobre as realizações e sem concluir sobre a eficácia ou eficiência das intervenções. Em cerca de um

quarto dos casos lista-se apenas o conjunto de atividades conduzidas, sem qualquer informação indicativa do seu sucesso ou insucesso, nem apreciação crítica, como no exemplo abaixo:

Estrutura Residencial para Idosos

No ano de 2013 esta Estrutura Residencial viu a sua capacidade alargada para 22 clientes. Ao longo deste ano realizaram-se diversas atividades de estimulação e de animação com a finalidade de promover a autoestima e a habilidade física e mental dos clientes, maioritariamente dependentes e sem capacidade para a realização das suas atividades de vida diária.

Salientamos:

- Jogos lúdicos, ginástica geriátrica, exercícios de estímulo mental.
- Pintura e expressão plástica, uma das atividades mais apreciadas pelos clientes.
- Culto da religião e da fé, com o rezar diariamente do Terço à Virgem.
- Realização de convívios com as famílias, e com clientes de outros Equipamentos, em dias comemorativos.
- Comemoração do aniversário de cada cliente com a confeção de um bolo.

Fonte: Relatório de Atividades de OTS.

Mas se em muitos casos os elementos qualitativos e apreciativos apresentados se traduzem **em observações vagas e pouco rigorosas** sobre a forma como as atividades decorreram e sobre os resultados alcançados, em cerca de um quarto dos casos as análises críticas de caráter qualitativo são fundamentadas em observações mais sistemáticas e rigorosas, baseadas, designadamente, em análises globais de fichas de registo/ocorrências, de fichas de caracterização/diagnóstico/avaliação ou de planos de desenvolvimento individuais, em inquéritos de satisfação e em reuniões de avaliação com colaboradores, pessoas envolvidas nas atividades, equipas e serviços de terreno. Traz-se à colação um trecho do relatório de atividades de uma entidade em que é visível uma preocupação com a fundamentação das apreciações sobre o desempenho da organização, neste caso, a partir da informação recolhida em reuniões de avaliação e inquérito a equipas de intervenção:

O trabalho de proximidade, centrado efetivamente nas pessoas e nos seus contextos de vida é referido pelas equipas como o garante de um maior conhecimento das suas necessidades e de uma maior adequação das ações desenvolvidas. As pessoas envolvem-se mais na construção de soluções para os seus problemas, assumindo uma postura de maior responsabilização no seu processo de mudança. Contudo, continuam a verificar-se alguns obstáculos à participação, por exemplo, a precariedade económica não permite a deslocação das pessoas em transporte público, a reduzida oferta de educação e formação para jovens e adultos dificulta o seu envolvimento em percursos pessoais de aumento de qualificação o que, conseqüentemente, impede a integração qualificada no mercado de trabalho. (Relatório de atividades de ADL - 7)

Os **exercícios mais estruturados de cariz quantitativo**, por seu turno, surgem mais frequentemente enquadrados em **processos de gestão da qualidade** (certificação), da

gestão estratégica (planeamento estratégico) e da **gestão do desempenho** (gestão centrada nos resultados). Nestes casos, as práticas que as organizações identificam como práticas de avaliação assumem as características dos processos de **medição do desempenho**, que tendem a enfatizar **indicadores quantitativos** com vista a apreciar a desempenho face aos resultados esperados. Todos estes processos obedecem à **lógica da gestão por objetivos** (típica da avaliação de desempenho na administração pública), enquanto metodologia norteada pelo estabelecimento de metas de trabalho, tendo em conta os recursos disponíveis e o período de tempo previsto para a sua realização e que assenta, basicamente, na adoção de um sistema de gestão radicado num conjunto de indicadores, tendo como prioridade o cumprimento de metas pré-definidas (em geral, quantitativas). A título ilustrativo, apresenta-se de seguida excerto de quadro de balanço anual apresentado em relatório de atividades de organização entrevistada.

Quadro 11. Balanço anual de atividades de OTS

Objetivos	Ações	Meta prevista	Indicadores de acompanhamento	Resultados	Análise do desvio e ações realizadas
Promover a responsabilidade ambiental	Sensibilizar para uma utilização responsável dos recursos	>=98,5%	Taxa de execução orçamental	A taxa de execução orçamental para 2012 foi de 99,3%.	Verificou-se a superação de objetivo.
Promover a responsabilidade social	Sensibilização para a problemática da deficiência e incapacidade	>= 12	Nº de ações	No final de 2012 foram desenvolvidas 27 ações	Verificou-se a superação do objetivo.

Fonte: Relatório de atividades de IPSS – 8.

Quando questionadas sobre que tipo de informação costuma recolher com maior frequência as respostas são equilibradas: **em metade das entidades é privilegiada a informação quantitativa e na outra metade, qualitativa**. Apesar do peso equilibrado da informação qualitativa e quantitativa veiculado nas respostas a esta questão, a especificação do **tipo de informação recolhida aponta para a prevalência da informação de carácter quantitativo**. A informação de avaliação mais comum é, pois, de natureza quantitativa e respeita às despesas e às realizações (n.º de ações e pessoas beneficiárias): mais de dois terços das organizações afirmam recolher este tipo de informação com base

regular (respetivamente, 83%, 81% e 68% das organizações). Elementos sobre as mudanças nos comportamentos, atitudes ou situações suscitadas pelas intervenções, e que são tendencialmente qualitativas, são recolhidos com menor frequência: menos de metade das organizações (44%) recolhe informação sobre o impacto regularmente.

As entrevistas também apontam para a prevalência de informação quantitativa, sobretudo sobre elementos de execução financeira e física de projetos ou serviços, mas também emergentes dos inquéritos de satisfação, cuja análise é, em geral, extensiva, muito orientada para o cálculo de taxas de satisfação com os produtos ou serviços prestados e menos para as *nuances* da experiência e opiniões das pessoas. A informação sobre a satisfação das pessoas beneficiárias, recolhida sobretudo através de questionários, é recolhida com regularidade por mais de metade das organizações inquiridas (53%).

Os métodos e técnicas de avaliação mobilizados para a recolha de informação reforçam a prevalência do “paradigma” quantitativo. Com efeito, o uso de dados de fontes administrativas (fichas de registo de atividades, etc.), em geral quantitativos, recolhidos de forma rotineira como parte integrante da gestão de projetos ou serviços, é a situação mais frequente nas organizações estudadas: 95% das inquiridas e 100% das entrevistadas recolhem este tipo de informação de avaliação. Os dados administrativos são gerados principalmente durante o processo de implementação das atividades e através da monitorização contínua. **A recolha de dados preexistentes** (dados administrativos e outros dados documentais), sendo menos exigente em termos de recursos técnicos e económicos, **é particularmente acessível à generalidade das organizações.** O mesmo não parece acontecer com a generalidade das peças metodológicas associadas a abordagens qualitativas, sobretudo *focus group* e estudos de caso, com uma adesão relativamente modesta das organizações abrangidas no estudo (respetivamente, de 25% e 21% das entidades inquiridas), mas também entrevistas individuais e observação direta (a que aderem 40% e 41% de entidades inquiridas). Além de pouco frequentes, como se confirma (e acentua) nas entrevistas em profundidade, a utilização destes métodos (sobretudo enquadrando abordagens multi-método) – e de outras técnicas inovadoras participativas (visuais, escritas e orais) –, **associa-se especialmente à participação em projetos.** Um outro instrumento de monitorização e avaliação muito referenciado no estudo, que podemos

associar a abordagens qualitativas, e que constituem modalidades mais ou menos informais de avaliação, são as reuniões de equipa. Embora estes encontros não possam ser sempre identificados como momentos formais de avaliação, ou porque não são conduzidos especificamente com esse objetivo, ou porque não há sistematização da informação recolhida, eles são entendidos como momentos estratégicos de reflexão sobre as formas como as atividades são implementadas e sobre os seus resultados, fazendo emergir sobretudo informação qualitativa sobre o desempenho. Os excertos que se seguem espelham bem a relevância destes momentos:

A organização até faz muito e perde muito tempo com isso. Somos bastante exigentes, mas é este tipo de avaliação que eu lhe digo: fazemos muitas reuniões de nos sentarmos à mesa e discutir, isso é que é avaliação! (...) Não é registado. É uma ata, não é registado em termos do processo avaliativo, mas é esta avaliação que fazemos. (Avaliação com técnico de ONG - 14)

Aí é uma lacuna que nós temos. Ou seja, trabalhamos ainda muito de forma empírica, não há uma sistematização da informação nem a trabalhamos com base numa metodologia específica. (Avaliação com dirigente de ONG - 34)

As análises estatísticas realizadas a partir do inquérito às organizações revelam que, em termos globais, a **informação quantitativa preexistente** (especificamente dados administrativos relativos a despesas e realizações), sendo a mais frequentemente recolhida nas organizações, **é também a menos influenciada por características organizacionais**. Vemos que estas são particularmente relevantes para a compreensão das diferenças na utilização de alguns métodos qualitativos de recolha direta de dados, especialmente a entrevista individual e os grupos de trabalho de avaliação, mas também de um método frequentemente associado a abordagens quantitativas – o inquérito (a que aderem metade das organizações) –, e sobre cujas distribuições exercem influência estatisticamente significativa todas as características organizacionais.

Apresenta-se de seguida um quadro com o perfil das organizações associadas à maior ou menor adesão a cada um dos suportes da recolha de informação e respetivos *p-values* (entre parêntesis), estimados a partir do teste de Qui-quadrado (consideram-se estatisticamente significativos os efeitos cujo *p-value* é inferior ou igual a 0,05).

Quadro 12. Suportes da recolha de informação: perfis de OTS

	DIMENSÃO	TIPO ENTIDADE	RECEITAS	FUNÇÕES	IDADE
Fontes secundárias					
Dados administrativos (registos de atividades, de participantes, etc.)	Médias/Grandes (0,002)		Públicas/Próprias (0,050)		
Despesas	> (0,034)				
N.º de atividades			Públicas/Filantropia (0,015)		< (0,013)
Outros dados documentais (relatórios, atas, etc.)		IPSS/ADL/ONG (0,043)	Públicas/próprias (0,015)	Serviço/Mistas (0,003)	
Fontes primárias					
Inquérito	> (0,000)	ADL → IPSS → ONG (0,000)	Públicas (projetos) (0,000)	Serviço → Mistas (0,000)	Recente/Intermédia (0,000)
GT sobre avaliação	Grandes (0,032)	IPSS/ONG (0,0000)	Públicas (subsídios)/Filantropia (0000)	Serviço → Mistas (0,000)	Recente/Intermédia (0,000)
Entrevistas individuais	> (0,001)	IPSS (0,000)	Públicas (subsídios) / Filantropia (0,001)	Serviços → Mistas (0,000)	< (0,011)
Focus Groups		ONG (0,061)*	Filantropia (0,000)	Mistas → Serviços (0,006)	Recente/Intermédia (0,022)
Observação estruturada	> (0,002)	IPSS → ADL (0,001)	Filantropia	Serviço/Mistas (0,000)	
Estudos de caso	<i>Sem diferenças estatisticamente significativas</i>				

Fonte: Inquérito por questionário.

Nota: *Diferenças estatisticamente pouco significativa.

Dentre as diversas características da estrutura das organizações, **o perfil da atividade das organizações é particularmente relevante para a compreensão das diferenças na utilização da maior parte dos métodos de recolha direta de dados**, especificamente do inquérito, dos grupos de trabalho sobre avaliação, da entrevista individual e da observação estruturada. Aqui, as diferenças entre organizações com funções de serviço (sobretudo IPSS e ADL) e entidades que desenvolvem sobretudo atividades de expressão (particularmente as coletividades), são fortemente marcadas: nas

primeiras a utilização das fontes de informação primária é bastante mais frequente do que nas segundas.

Além do perfil da atividade, a análise ressalta também a influência significativa do **perfil de receitas** na seleção de fontes de informação, quer primária, quer secundária. É interessante verificar que as organizações cujos fundos advêm da filantropia (doações individuais, doações de fundações e doações empresariais) são as que menos se associam à recolha de dados preexistentes (designadamente administrativos) e às fontes primárias associadas a análises extensivas/quantitativas (inquéritos, por exemplo), aderindo particularmente a técnicas de recolha de informação primária associadas a análises intensivas/qualitativas (como sejam as *focus group*). Por outro lado, a **menor dimensão financeira** prenuncia o recurso significativamente inferior à generalidade das técnicas de recolha de informação, à exceção de algumas associadas a análises intensivas (*focus group*, estudos de caso e análises de documentos).

A associação das organizações de pequena dimensão ao paradigma qualitativo, que parecem indicar estes resultados, é reiterada pelas análises realizadas a partir da questão colocada às organizações sobre a natureza da informação mais frequentemente recolhida. Com efeito, verificou-se que enquanto a informação quantitativa é significativamente mais representativa das organizações de grande dimensão, dependentes de receitas públicas (provenientes de projetos), com funções de serviço, a informação qualitativa é predominante nas organizações pequenas, com funções mistas ou de expressão e autossustentadas¹⁰¹. No entender de alguns autores (e.g., Hall, 2014; Davies, 1998), algumas das principais abordagens/técnicas qualitativas de avaliação foram desenvolvidas nas e para organizações do terceiro setor e por isso aparecem mais cuidadosamente harmonizadas com as exigências que as técnicas de avaliação podem colocar em termos de recursos e competências. Tais abordagens, entre as quais se destaca a das “Mudanças Mais Significativas” (MSC)¹⁰², **alinhando-se com uma lógica de avaliação**

¹⁰¹ Para avaliar se a natureza da informação recolhida varia segundo as principais organizacionais recorreu-se ao Teste de Qui-quadrado seguindo os procedimentos e condições já especificados. Concluímos que a recolha de dados quantitativos/qualitativos é significativamente influenciada pelo perfil de receitas ($p=0,000$), dimensão ($p=0,001$) e pelo perfil da atividade ($p=0,001$).

¹⁰² A técnica das “Mudanças Mais Significativas” foi desenvolvida na década de 1990, enquanto abordagem para avaliação de programas de desenvolvimento social complexos. No centro da técnica estão

centrada na aprendizagem, são eventualmente mais vantajosas em termos de custos do que as abordagens quantitativas desenhadas para satisfazer o ideal de “prova”, sob uma lógica científica de avaliação, como é o caso da metodologia SROI e do quadro lógico¹⁰³ (Hall, 2014). Sendo menos exigentes em termos de custos e competências, algumas abordagens qualitativas são mais acessíveis às organizações de pequena dimensão financeira, nos quais o financiamento para a avaliação é limitado ou inexistente. Por outro lado, a exigência de informação quantificada sobre o desempenho está particularmente associada a técnicas desenvolvidas por entidades financiadoras e doadoras (como é o caso do SROI e do quadro lógico), num cenário em que o dinheiro e experiência necessários estão eventualmente mais disponíveis e acessíveis (*ibidem*: 332).

Todavia, a análise feita das entrevistas e dos documentos revela que, frequentemente, as abordagens que técnicos/as e dirigentes das organizações identificam como qualitativas consistem em apreciações rápidas e intuitivas sobre a forma como as atividades decorreram e os resultados que foram alcançados, não tendo por base procedimentos sistematizados de recolha e análise de informação. Como vimos atrás, é sobretudo a partir deste tipo de abordagem qualitativa de carácter pouco rigoroso e sistemático que se delineiam elementos sobre os impactos das intervenções. Este parece ser o caso de algumas IPSS de pequena ou média dimensão auscultadas, em cujos relatórios de avaliação (que são geralmente os relatórios de atividades anuais) a informação qualitativa sobre os efeitos da sua intervenção nos utentes resulta de impressões vagas do pessoal técnico a partir da sua relação de proximidade com as pessoas que servem e do *feedback* que delas vão recebendo.

a recolha e interpretação sistemáticas de histórias sobre mudanças importantes suscitadas por determinado projeto, as quais são tipicamente preparadas por pessoas mais diretamente envolvidas de forma regular nas iniciativas, tais como beneficiários, clientes e pessoal do terreno. Para mais informação sobre a abordagem, consulte-se <http://mande.co.uk/special-issues/most-significant-change-msc/>, and www.clearhorizon.com.au/flagship-techniques/most-significant-change/.

¹⁰³ O quadro lógico/analítico (*logical framework*), enquadramento lógico, foi desenvolvido ao longo dos anos de 1970 enquanto ferramenta de planeamento e avaliação, direcionada sobretudo para grandes organizações doadoras bilaterais e multilaterais, tais como a USAID, o Departamento de Desenvolvimento Internacional (Reino Unido), o Programa das Nações Unidas para o Desenvolvimento e Comissão Europeia. A técnica disseminou-se por muitas organizações do terceiro setor, à medida que estas foram beneficiando crescentemente de fundos destas e outras organizações doadoras que utilizam esta metodologia. Trata-se de uma ferramenta de estruturação da lógica de uma intervenção, que assenta numa apresentação matricial de realizações, resultados e efeitos para cada um dos níveis de objetivos.

Síntese

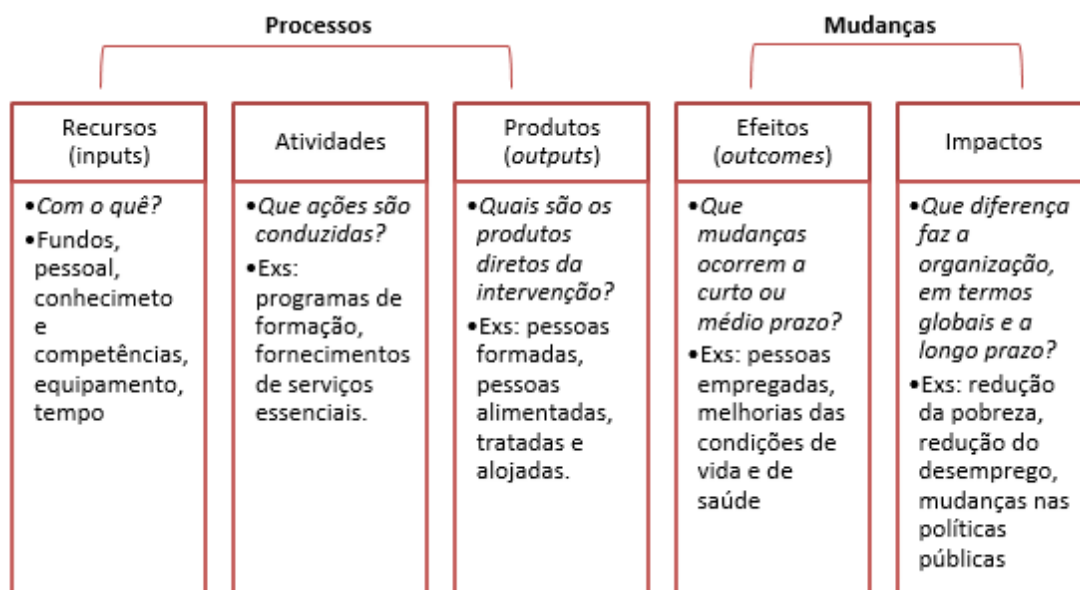
Ao mesmo tempo que se reconhece a pertinência das abordagens qualitativas, até nas orientações metodológicas de proeminentes agências internacionais (e.g., Banco Mundial e Comissão Europeia), na prática privilegiam-se as lógicas quantitativas da medição, sobrepondo-se a importância da quantificação das realizações e dos resultados, mediante o acompanhamento de indicadores de desempenho (que medem os resultados face aos objetivos preestabelecidos). O paradigma quantitativo/descritivo encontra-se vinculado às abordagens mais formais e sistemáticas, e está particularmente associado à prestação de contas ascendente, apresentando um carácter mais coercivo comparativamente à recolha de informação de natureza qualitativa. O paradigma qualitativo/crítico, por seu turno, tendendo a apresentar um carácter voluntário, encontra-se sobretudo vinculado a abordagens informais, assentes em apreciações rápidas e intuitivas sobre a forma como as atividades decorreram e os resultados que foram alcançados, sem recurso a métodos formais de produção de informação. Por outro lado, a forte tendência da generalidade das organizações em quantificar a sua intervenção ocorre essencialmente ao nível das realizações/*produtos/outputs* das intervenções. Os elementos sobre as mudanças de mais longo prazo das intervenções tendem a ser delineados a partir de informação qualitativa assente em impressões do pessoal técnico e/ou fontes informais. Ou seja, a afirmação do paradigma qualitativo nas organizações estudadas faz-se essencialmente a partir deste tipo de abordagem informal, de carácter pouco rigoroso e sistemático. A mensuração de alguns efeitos “objetivos” mais facilmente quantificáveis, por seu turno, implicaria a mobilização de recursos para auscultar as populações decorrido determinado período após o fim da intervenção que muitas organizações não têm condições e/ou estão disponíveis para fazer. Na verdade, os desafios que se colocam aos processos de avaliação em termos de conceção e execução são ampliados à medida que uma organização se move de uma avaliação de atividades e *outputs* para uma avaliação de efeitos (*outcomes*) ou impactos, sobretudo mediante abordagens quantitativas construídas sob o ideal de prova. O carácter imaterial e subjetivo de muitas das mudanças suscitadas pelas intervenções sociais desafia – quando não impossibilita – a sua quantificação.

4.4.2. Âmbitos de incidência e conteúdos da avaliação: o enfoque nos processos e realizações e o desafio da orientação para os impactos

Desde a década de 1990, em muitas economias ocidentais, especialmente no designado mundo anglo-saxónico, tem sido crescente o movimento no sentido de tratar os resultados e impactos como uma parte fundamental de medição de desempenho no setor público. No Reino Unido, por exemplo, as metas departamentais – anteriormente baseadas em *outputs* – passaram a estar cada vez mais focadas nas mudanças suscitadas pela intervenção dos serviços (Ellis e Gregory, 2008: 3). Em Portugal, ao nível do setor público, o Sistema Integrado de Gestão e Avaliação do Desempenho na Administração Pública (SIADAP) – no enquadramento da Nova Gestão Pública da Gestão Orientada para os Resultados – é ilustrativo da deslocação do enfoque dos recursos e atividades para os resultados imediatos das intervenções – os *outputs*. **A transição para os resultados de mais longo prazo não parece ser tão visível**, segundo é possível observar a partir da análise dos objetivos operacionais e indicadores integrados nos Quadros de Avaliação e Responsabilização (QUAR) de alguns organismos do Estado¹⁰⁴ (vide figura 5 com cadeia de resultados das intervenções).

¹⁰⁴ Foram selecionados e analisados documentos QUAR relativos a 2015 de 10 organismos: Instituto da Segurança Social, IP, Direção Geral de Agricultura e Desenvolvimento Rural, Caixa geral de Aposentações, Programa de Desenvolvimento Rural, Direção-Geral do Ensino Superior, IAPMEI - Agência para a Competitividade e Inovação, I.P., Instituto Nacional de Emergência Médica, Secretaria-Geral do Ministério da Saúde, Secretaria-Geral do Ministério da Justiça e Direção-Geral de Saúde. É evidente a centralidade dos **indicadores de processo** (por exemplo, n.º de pareceres emitidos sobre diplomas legais, n.º de contributos em grupos de trabalho, % de respostas sociais com acordo de cooperação acompanhadas, grau de satisfação dos utilizadores), com o intuito de avaliar a eficácia, eficiência e a qualidade da gestão pública e a ausência de indicadores de impacto que traduzam os benefícios ou as consequências da intervenção pública nas pessoas destinatárias ou na população mais vasta.

Figura 5. Cadeia de valor da intervenção das organizações



Fonte: Elaboração da autora.

Em todo o caso, as **transições** que se verificam ao nível da medição do desempenho no setor público são **relevantes para as organizações do terceiro setor**, por **via da relação que estabelecem com o Estado para a concretização dos seus objetivos em domínios diversos** (saúde, educação, ação social, desenvolvimento local, etc.), mediante a realização de acordos de cooperação ou do desenvolvimento de projetos financiados por fundos públicos.

A centralidade de **indicadores financeiros, de realização e de resultados imediatos, facilmente mensuráveis**, é transposta para as relações do Estado com o Terceiro Setor, em especial no contexto de contratualização de serviços de apoio social com as Instituições Particulares de Solidariedade Social (IPSS). A informação requerida no contexto destas relações incide, basicamente, na **recolha de indicadores de realização física e financeira**¹⁰⁵, sem preocupação com efeitos da ação das IPSS nas pessoas

¹⁰⁵ A apresentação de contas e orçamentos por parte das IPSS é um conjunto de obrigações que estas passam a ter que cumprir a partir do momento em que se registam como IPSS na Segurança Social, assim como o envio regular de listas de frequência mensal de utentes, no caso das organizações com acordos de cooperação. A generalidade das entidades auscultadas, para além dos relatórios de contas e orçamentos, envia também anualmente à tutela os relatórios e planos de atividade, sem que estes obedeçam a uma estrutura específica, como acontece com os de contas (obedecem a um modelo definido pelo ISS). Contudo, como não recebem qualquer *feedback* dos relatórios de atividade enviados, a conceção que prevalece é a de

beneficiárias das respostas sociais ou nos problemas sobre os quais procuram intervir. Estes elementos configuram um **tipo de avaliação de controlo financeiro**, decorrente do conjunto de obrigações contabilísticas que as IPSS têm que cumprir e das participações que recebem no âmbito dos acordos de cooperação com a Segurança Social. Além da recolha deste tipo de informação – da responsabilidade das instituições – os Centros Distritais de Segurança Social fazem visitas de acompanhamento às instituições e respostas sociais com o objetivo de avaliar o cumprimento das disposições legais e dos normativos específicos de cada resposta social, sendo dada especial atenção ao cumprimento da legislação referente a “aspectos específicos das instituições, normas relativas ao edificado e normas específicas das respostas sociais” (ISS,IP, 2007: 24). O enfoque não é, pois, **nos benefícios resultantes da intervenção das instituições** para as pessoas beneficiárias, mas sim no cumprimento dos requisitos legais de funcionamento, numa lógica que se aproxima do conceito de **auditoria**.

Mas o enfoque dos requisitos de prestação de contas nos **elementos técnicos e financeiros da implementação** das atividades não parece ser um exclusivo da relação de cooperação para a prestação de serviços sociais, entre Estado e IPSS. Uma crítica reiteradamente formulada em contexto de entrevista com responsáveis de OTS prende-se exatamente com o **enfoque excessivo das entidades financiadoras nos aspetos financeiros e técnicos associados à implementação das intervenções, sem a atenção desejável em relação às mudanças suscitadas pela ação das organizações**. O testemunho que se segue é ilustrativo deste tipo de apreciação:

As entidades, seja o governo português ou outras, financiam coisas que não sabem o que estão a financiar, e não sabem o resultado, o impacto ou que avaliação existe. Na realidade, isso não interessa. Ele até pode estar lá, porque todas as candidaturas e todos processos implicam que haja, há sempre a questão “como é que vocês avaliam?”. Mas depois na prática o que me parece é que isso não tem importância. (Entrevista com dirigente de ONG – 28)

Esta crítica é extensível a boa parte das intervenções financiadas ou cofinanciadas pelo Estado no âmbito de programas nacionais ou comunitários, mas também a alguns financiamentos internacionais, designadamente da Comissão Europeia.

que eles nem sequer são analisados pela tutela, que apenas considera os elementos obrigatórios (dados financeiros e registos de frequências mensais de utentes).

Um dos programas selecionado para análise mais aprofundada – o POPH – é identificado como arquétipo de uma perspectiva de avaliação centrada nos **recursos, nas atividades e nas realizações físicas e financeiras**. Este enfoque espelha uma visão restrita da avaliação, como comenta uma pessoa entrevistada ligada à coordenação do QREN:

Os requisitos são relativamente *soft*, em muitos casos prevaleceu até uma perspectiva um bocadinho débil do próprio conceito de avaliação: confundido com a apresentação de evidências de resultados e realizações. (Entrevista com coordenação do QREN - 1)

O que é pedido é que se insira no sistema de informação uma série de elementos de execução que incidem sobretudo sobre os processos (e muito pouco sobre os efeitos das intervenções). Alguma da informação requerida serve para alimentar os indicadores de desempenho (de realização e de resultado) definidos ao nível de cada eixo prioritário do Programa. Todavia, **nenhum dos indicadores definidos para o próprio Programa são indicadores de impacto**, pelo que a questão da ausência de análises de impacto coloca-se a montante dos próprios projetos, como reconhece responsável da gestão do POPH em entrevista:

As atividades de avaliação dessas atividades ou dos impactos (...) nunca foi bem resolvida nestas questões dos fundos. Admito que o que tenha sido melhor tenha sido o EQUAL que no final ainda conseguiu. Mas admito que é uma questão sempre mal resolvida. Até porque para aferir convenientemente os impactos de qualquer projeto ou atividade é preciso algum distanciamento temporal. E esse distanciamento temporal nem sempre é compatível com os procedimentos, financiamentos e prestações de contas que estão associados a um QCA ou estes programas que têm prazos para prestação de contas. (Entrevista com responsável da gestão do POPH - 3)

Aliás, os indicadores de realização e de resultado¹⁰⁶ definidos para o programa são de natureza muito semelhante (por vezes difíceis de distinguir entre si), remetendo mais para atividades e para produtos das intervenções do que para os seus efeitos e benefícios sobre as pessoas beneficiárias/as ou população mais vasta. Apresentam-se de seguida, a título de exemplo, os indicadores definidos para os eixos prioritários 2 “Adaptabilidade e Aprendizagem ao Longo da Vida” e 7 “Igualdade de Género” do Programa.

¹⁰⁶ No QREN (Quadro de Referência Nacional), em que se enquadra o POPH, os indicadores de realização (output) medem “o produto material, ou output, gerado diretamente pela atividade da intervenção, podendo ser expresso em unidades físicas ou monetárias. Os indicadores de resultado, por seu turno, medem “o efeito direto e imediato do produto gerado por uma intervenção sobre os seus beneficiários, podendo ter um carácter material ou imaterial.” (Fonte: <http://www.qren.pt/np4/3085.html>)

Quadro 13. Indicadores de realização (output) e resultado do POPH

	Indicadores de Realização (output)	Indicadores de Resultado
Eixo Prioritário 2	<ul style="list-style-type: none"> – N.º de Centros/Equipas Novas Oportunidades Apoiados – N.º de Abrangidos em Centros Novas Oportunidades – N.º de adultos abrangidos em cursos de educação formação – N.º de adultos abrangidos em formações modulares certificadas 	<ul style="list-style-type: none"> – N.º de adultos que obtiveram dupla certificação – N.º de adultos que obtiveram uma certificação escolar ou profissional – N.º de adultos que obtiveram competências escolares e/ou profissionais certificadas
Eixo Prioritário 7	<ul style="list-style-type: none"> – N.º de projetos promovidos por ONG/instituições sem fins lucrativos apoiados – N.º de empresas apoiadas para desenvolverem planos para a igualdade – N.º de pessoas abrangidas em ações de formação em igualdade de género – N.º de pessoas abrangidas em ações de formação em violência de género – N.º de mulheres abrangidas em ações de consultoria e assistência técnica à criação e desenvolvimento de microempresas geridas por mulheres – N.º de ações de Informação e Divulgação 	<ul style="list-style-type: none"> – N.º de pessoas abrangidas pelos projetos promovidos pelas ONG/instituições sem fins lucrativos – N.º de trabalhadores e trabalhadoras envolvidos em planos empresariais para a igualdade – N.º de empresas criadas – N.º de pessoas abrangidas que concluíram ações de formação em igualdade de género – N.º de ONG apoiadas com projetos de promoção da igualdade de género – N.º de pessoas abrangidas que concluíram ações de formação em violência de género

Fonte: Documento de programação do POPH.

Além da informação que responde diretamente a estes indicadores, as entidades promotoras dos projetos devem reportar um volume significativo de dados sobre a implementação/execução física das ações – caracterização das atividades (caracterização, destinatários/as, situação atual, etc.) e pessoas destinatárias/entidades apoiadas (elementos de identificação, número, tipo, etc.) –, e sobre a execução financeira dos projetos (declaração de despesas).

No entender da generalidade das organizações entrevistadas que estão ou já estiveram envolvidas em projetos POPH, este tipo de informação é **desajustada quer para aferir sobre a qualidade dos projetos financiados quer para melhorar as intervenções**, servindo apenas o propósito de **prestar contas sobre a concretização das atividades** e do orçamento previstos em candidatura, numa espécie de “avaliação automática” pela plataforma do SIIFSE, como qualifica responsável de organização de cúpula em entrevista. Os testemunhos que se reproduzem de seguida, não só de promotores de projetos mas

também da gestão do programa, são ilustrativos da reduzida atenção do programa aos impactos das ações conduzidas pelos projetos:

O registo no próprio sistema de informação de alguns indicadores. Esses indicadores dão para retirar algumas conclusões mais de uma perspetiva de realização física do projeto do que numa perspetiva de avaliação dos impactos, ou noutras perspetivas. (Entrevista com gestão do POPH - 3)

Tem a ver com alguns requisitos, nomeadamente o preenchimento da execução física que é necessário preencher (pessoas, sexo, qualificação). Mas o que nós queremos perceber é alguma coisa mais: perceber se teve ou não alcance, como é que teve ou não conhecimento, para depois fazermos melhor. (Entrevista com dirigente de ONG – 27)

Exige no sistema *online*, a meter lá os resultados. E portanto toda a parte técnica é quantificada *online*. Mas a parte mais de impacto não é exigido. (...) Essas entidades interessam-se mais pelos resultados do que pelo impacto. (Entrevista com dirigente de OTS – 9)

Sobre os requisitos de avaliação no âmbito do **PIC EQUAL**, por seu turno, a análise dos documentos de programação e de outra documentação produzida no âmbito do programa e as perspetivas das organizações beneficiárias colhidas a partir do inquérito por questionário e das entrevistas, apontam para uma **maior atenção em relação aos efeitos de médio e longo prazo das intervenções financiadas**.

Em todos os regulamentos, guias de apoio à conceção, implementação e avaliação de projetos editados pelo PIC EQUAL aparecem referências à necessidade de fazer incidir os exercícios de avaliação não apenas nos planos da conceção (diagnóstico, destinatários, objetivos, plano de ação, parcerias, ...), da operacionalização (fluxos financeiros, de informação, organização da gestão/dos projetos, ...) e realização (desempenho, grau de execução física e financeira), mas também dos **efeitos/impactos** (contextos, atores, sustentabilidade, ...). No guia de autoavaliação do PIC EQUAL pode ler-se:

A avaliação no contexto de um Programa de Intervenção como é a Iniciativa Comunitária EQUAL assume um papel central na compreensão da própria intervenção ao nível de cada Projeto, permitindo recolher, sistematizar e analisar informação sobre o modo como aquela se está a desenvolver e os resultados e impactes que pode produzir. (Gabinete de Gestão EQUAL, 2003: 6).

No plano dos impactos, convidam-se os projetos a aferir os efeitos das ações empreendidas, “tanto os diretos como os indiretos, os desejados e os não desejados”, distinguindo-se, nomeadamente, “os efeitos produzidos no contexto de partida (aferidos a

partir do diagnóstico e dos Balanços de Competências)¹⁰⁷ e os efeitos de disseminação e *mainstreaming* de práticas e produtos inovadores” (Gabinete de Gestão EQUAL, 2003: 13). O Balanço de Competências é identificado como sendo o instrumento privilegiado de verificação de parte dos impactos dos projetos ao nível dos recursos humanos; enquanto os outros tipos de efeitos são monitorizados e avaliados através da autoavaliação e da avaliação independente. Em ambos os exercícios é vincada a importância de atender aos efeitos de curto, médio e longo prazo das intervenções. Nas “orientações para a avaliação independente dos projetos EQUAL”, por exemplo, exortam-se as entidades responsáveis pela avaliação, em sede de avaliação no final do projeto, a dar resposta às componentes e questões de avaliação, “com especial incidência na identificação e análise dos resultados e potenciais impactos da intervenção” (Gabinete de Gestão EQUAL, 2005: 3). Da parte das organizações envolvidas em projetos EQUAL entrevistadas, existe um claro reconhecimento da orientação para o impacto da iniciativa, tanto ao das pessoas e entidades destinatárias, como das pessoas e organizações agentes das próprias intervenções:

Mas eu acho que enquanto ao nível do EQUAL havia uma preocupação muito grande que este tipo de projetos tivesse impacto nos beneficiários, mas também ao nível da organização, e que existisse um conjunto de indicadores que medissem a avaliação do impacto ao nível da organização. Pelo menos a organização acabava por ser impelida a fazer esse exercício de autoavaliação, acho que aqui [POPH] não. Aqui respondemos a um conjunto de critérios que o financiador entende que são importantes para eles, mas que nós não sentimos que sejam assim tão decisivos. (Entrevista com dirigente de ADL – 20)

Apesar da **exigência em termos de avaliação de impacto ao nível dos projetos financiados por fundos públicos ser ainda limitada e por isso pouco presente nas preocupações da coordenação dos projetos**, as experiências relatadas nas entrevistas sobre a condução de **avaliações de impacto têm sobretudo lugar no âmbito de projetos específicos**, seja financiados pelo Estado (particularmente no âmbito do PIC EQUAL, do PRODER e da Cooperação para o Desenvolvimento)¹⁰⁸, seja fora do quadro da relação com

¹⁰⁷ O Balanço de Competências é definido no âmbito da IC EQUAL como “um instrumento de apoio à gestão, com funções de diagnóstico e de avaliação das competências mais diretamente relacionadas – mobilizadas ou desenvolvidas - com os objetivos e atividades de um determinado projeto. É também um meio de demonstrar objetivamente as aquisições, em termos de competências, do projeto” (Estevão, 2005: 7).

¹⁰⁸ Estas linhas de financiamento obrigam à apresentação de relatórios de avaliação final e/ou intercalar, disponibilizando verbas para a realização de estudos de avaliação.

o Estado, tais como fundações filantrópicas e empresariais, ou mesmo no âmbito de atividades de investigação no enquadramento de estágios académicos ou de parcerias com instituições universitárias:

Fizemos isso sobretudo em termos de projetos específicos. Um [dos trabalhos] foi perceber o que é que as pessoas que nós temos ajudado a conseguir emprego, qual era a sua situação passado 15 anos. (...) Isso foi um estudo que está publicado, que foi a colega que fez no âmbito da sua tese de mestrado. Depois houve um estudo pelo X no âmbito de outro projeto que fizemos em termos de algumas medidas de impacto. Ou seja, essa medida de impacto não tem sido um processo muito continuado, o que tem a ver com recursos. (Entrevista com dirigente de ONG – 16)

Mas, na maioria destes casos, a avaliação é feita no final dos projetos (por imposição do financiamento), quando ainda não é possível identificar muitas das mudanças induzidas pelas intervenções, inviabilizando análises mais robustas sobre os efeitos de médio ou longo prazo:

Fazemos avaliação de impacto no final do período dos projetos, mas é uma avaliação de impacto à boca do fim do projeto. E depois não conseguimos qualquer financiamento para conseguir fazer uma avaliação do impacto do projeto passado 1 ano. É completamente diferente nós avaliarmos o impacto de um projeto de 3 anos, onde já se consegue no final do projeto avaliar alguns impactos, nomeadamente porque há algumas atividades e alguns resultados já concluídos há 6 meses. (Entrevista com dirigente de ONG – 6)

O foco nos impactos, **ainda com pouca tradução nas práticas das organizações do terceiro setor em Portugal, tem vindo a ser privilegiado, nos últimos anos, pelos fluxos de financiamento provenientes da filantropia**, já preocupados com a demonstração do impacto social das intervenções apoiadas, especialmente através de abordagens *value for money*. Exemplos desta abordagem podem ser encontrados nos requisitos dos Programas “Cidadania Ativa”¹⁰⁹, gerido pela Fundação Calouste Gulbenkian, e “EDP Solidária”¹¹⁰,

¹⁰⁹ Entre os critérios de seleção dos projetos encontram-se os mecanismos propostos para supervisão e avaliação e o retorno sobre o investimento (*value for money*). Além disso, conforme consta no Regulamento do Programa de 2014, os promotores de grandes projetos devem apresentar um relatório de seguimento sobre o impacto do projeto no prazo de seis meses após a conclusão do projeto, de acordo com o modelo disponibilizado pela unidade de gestão do programa (disponível em http://www.gulbenkian.pt/mediaRep/gulbenkian//files/institucional/actividades/programas_projectos/EEA_grants_CidAtiva/Docs/Regulamento_do_Programa_2014_PT.pdf).

¹¹⁰ Na apreciação dos projetos apresentados pelas entidades candidatas, a medição do impacto, mediante a incorporação de processos de monitorização e avaliação proporcionais à dimensão dos projetos, é um dos critérios considerados, como se lê no Regulamento do Programa “EDP SOLIDÁRIA 2014”(disponível em http://www.fundacaoedp.pt/folder/pagina/ficheiro1/437_Regulamento%20EDP%20Solid%C3%A1ria%202014%20%285.2.2014%29.pdf).

promovido pela Fundação EDP. Dentre este tipo de abordagem, **o SROI parece estar a assumir um destaque importante (mas recente) no nosso país por parte da filantropia empresarial**, mediante o suporte à realização de formações e outras ações de capacitação para a utilização da metodologia.

A análise da documentação disponibilizada pelas organizações – em particular os relatórios e planos de atividades – permite perceber, por seu turno, o tipo de abordagem prevalecente ao **nível da organização como um todo** (e não de projetos ou iniciativas específicas). Na linha do que Ruela e Albuquerque (2016) observaram no seu estudo sobre *accountability* no TS, a grande maioria dos relatórios anuais de atividades caracteriza-se por um forte enfoque em informação sobre a operacionalização a implementação das atividades, frequentemente com grande detalhe sobre os procedimentos adotados, mas sem informação sistemática sobre os seus **efeitos nas pessoas beneficiárias e/ou na comunidade**.

As mudanças de médio ou longo prazo produzidas pelas organizações **são descritas e analisados com pouca frequência e, em geral, de forma vaga e superficial**. Trata-se sobretudo de **análises críticas de carácter informal, resultantes de uma observação não sistemática** (por exemplo, a partir de impressões ou do *feedback* casual de pessoas beneficiárias) e centradas na visão do pessoal técnico envolvido nas das intervenções. Formulações vagas não fundamentadas como “as atividades permitiram o desenvolvimento pedagógico e a oportunidade de crescimentos social e integração na comunidade” (Relatório de atividades de ADL - 19), ou “as atividades participadas cumpriram a trave mestra da interculturalidade com a valorização da cultura e das tradições culturais (Relatório de atividades de ONG - 14), são frequentes quando as organizações se referem a mudanças de carácter qualitativo que são exigentes em termos de metodologias de verificação. O excerto de um relatório de atividades de uma organização abrangida pelo estudo, relativo à avaliação de uma atividade (estúdio de música), que tinha como objetivo o desenvolvimento de competências artísticas em jovens, ilustra a centralidade da perspetiva do pessoal técnico:

Considera-se que este trabalho teve um enorme sucesso, permitindo: o reconhecimento e valorização dos participantes; o aumento da afluência comunicativa; da autonomização; da responsabilização; da criatividade; da capacidade de integrar o seu sentir, o seu pensar e a sua ação; da capacidade de

liderança e partilha; o desenvolvimento das suas capacidades artísticas e pessoais bem como o melhoramento nas relações intrapessoais e interpessoais; o fortalecimento do “eu”, através da partilha de informações e técnicas. O envolvimento dos indivíduos constituiu um fator importante, conferindo-lhes um aumento do grau de motivação (que lhes permitiu decidir e incidir em temas do seu bem-estar), melhoramento do desempenho do projeto, bem como a contribuição na resolução e mediação de conflitos, gerando oportunidade de aprendizagem social. (Relatório de atividades de OTS)

Embora a **filosofia de gestão orientada para resultados** pareça permear os discursos e práticas de uma parte importante de organizações – e que é visível designadamente na adoção do planeamento estratégico e na avaliação do desempenho – **a conceção de resultados que prevalece vincula-se, pois, quase exclusivamente aos produtos diretos dos serviços ou das atividades das organizações** e, geralmente, são descritos ou medidos com base em indicadores de volume – frequência das ações e volume de participantes/taxas de execução. É muito menos frequente a conceção de resultados enquanto tradução dos efeitos das intervenções, mesmo dos efeitos que é possível identificar de forma direta e imediata junto das pessoas destinatárias.

Para além de análises de eficácia com base nos produtos das atividades, a nível organizacional também são frequentes as análises de satisfação de clientes, muito comuns nas organizações prestadoras de serviços. **A avaliação da satisfação, porém, traduz uma análise de processo e não de impacto**, por estar vinculada às atividades e por em geral não expressar as mudanças produzidas na vida das pessoas que beneficiam das intervenções.

A preocupação com a satisfação dos/as utentes, em especial, está muito associada aos sistemas de controlo da qualidade¹¹¹, os quais contêm um conjunto de linhas de orientação para implementar, na prática, os controlos de qualidade e a forma como essa qualidade deve ser medida e melhorada. Uma orientação transversal dos modelos de certificação de produtos/serviços (DGERT, ISO 9001, Modelo de Avaliação da Qualidade das Respostas Sociais – ISS e EQUASS) é exatamente a de **monitorização da satisfação das partes interessadas** – sobretudo clientes/utentes/utilizadores, mas também de trabalhadores/as, parceiros, financiadores e da comunidade –, o que explica a sua relevância para as organizações. Todavia, o enfoque deste tipo de avaliação é nos processos

¹¹¹ Controlo de qualidade é aqui entendido como “controlo sistemático da capacidade da organização para criar qualidade nos serviços que presta (SEAV, s/d).

conduzidos pelas organizações e não nas mudanças por eles suscitadas. Em todo o caso, é no enquadramento dos requisitos de **acreditação/certificação dos processos formativos, especificamente, que é produzida a maioria das referências à avaliação de impacto – avaliação de impacto da formação**. Aqueles requisitos obrigam as entidades, consoante a tipologia de projetos formativos desenvolvidos (destinatários, objetivos, duração), a recolher informação sobre a empregabilidade e inserção-profissional dos/as formandos/as:

Também fazemos avaliação de impacto, que é a integração dos formandos... feita passado 6 meses a um ano. (...) Para ver o grau de empregabilidade, se a formação foi bem-sucedida. Esta é uma avaliação que nós fazemos, que é extra qualquer situação de financiamento; tem a ver com a certificação da DGERT. (Entrevista com dirigente de Escola Profissional – 15)

Outra orientação dos sistemas de gestão da qualidade, em especial com base nos modelos EQUASS e do ISS, é a orientação para os resultados. No caso do modelo do ISS, o alcance dos resultados encontra-se sobretudo ligado à monitorização dos resultados de desempenho dos processos-chave¹¹² das respostas sociais à concretização dos objetivos individuais estabelecidos nos planos individuais dos utentes das respostas sociais, e à satisfação das partes interessadas (clientes, colaboradores/as e comunidade). Enquanto os indicadores de desempenho dos processos apresentados remetem essencialmente para a forma como os processos são conduzidos e não para os benefícios resultantes para as pessoas utentes (por exemplo, grau de cumprimento dos planos individuais, % de reclamações relativas ao processo de admissão e acolhimento, tempo médio de permanência em lista de candidato), **a monitorização e avaliação dos objetivos estabelecidos nos planos individuais**, dependendo do alcance dos objetivos que forem definidos no seio das entidades, **poderá ter um enorme potencial em termos de recolha de informação sobre as mudanças suscitadas pelas intervenções nas pessoas beneficiárias.**

No caso do modelo EQUASS, uma vez que certifica a entidade e não uma resposta social específica, a orientação para os resultados assenta também na **preocupação com a**

¹¹² Os processo-chave de cada uma das respostas sociais abrangidas pelo Programa de Cooperação para a Qualidade e Segurança das respostas Sociais encontram-se definidos nos manuais de processo-chave (Manuais de Gestão da Qualidade das Respostas Sociais) disponibilizados pelo ISS. Além dos processos específicos associados a cada uma das respostas sociais, incluem processos mais transversais como a candidatura, a admissão e acolhimento e o plano individual.

medição do desempenho global da organização face aos objetivos, além de medir os resultados alcançados e os benefícios para os clientes dos serviços recebidos, tanto a nível individual (a partir do plano individual), como coletivo. A inovação relativamente ao modelo do ISS reside também na distinção que se faz entre a medição dos resultados alcançados face aos objetivos estabelecidos em cada plano individual e a nível coletivo e os benefícios para as pessoas destinatárias em cada um dos planos, parecendo indicar uma **maior orientação para os impactos associados aos resultados das intervenções**. Esta orientação surge reforçada pelo enfoque na medição do valor acrescentado dos serviços prestados ao nível da qualidade de vida das pessoas servidas, que acrescenta à preocupação com satisfação. Apesar destas orientações, a análise dos documentos produzidos pelas organizações certificadas por este modelo, atesta a **dificuldade das organizações em orientarem os processos de monitorização e a avaliação para a análise dos benefícios que decorrem das atividades**, centrando-se nas atividades em si e respetivos *outputs*, mediante o acompanhamento de indicadores que medem a eficácia e a eficiência dos serviços. As análises de satisfação continuam a ser, na prática, a forma encontrada para aceder a alguma informação sobre a forma como os serviços prestados respondem às necessidades e expectativas das pessoas que deles beneficiam e das outras partes interessadas. Apresenta-se abaixo excerto de relatório de atividades de organização certificadas pelo modelo EQUASS, que é exemplar deste tipo de lógica de medição:

Objetivos	Ações	Meta prevista (Plano Atividades)	Indicadores acompanhamento	Resultados	Análise do desvio
Promover o sucesso contínuo do grau de execução dos Planos Individuais dos Clientes	Propostas de alteração/criação de instrumentos de avaliação e monitorização. Dinamização crescente de reuniões efetuadas pelo gestor de caso.	Grau de cumprimento dos Planos superior ao do ano anterior em pelo menos 1%	Número de instrumentos novos/alterados Grau de execução dos objetivos dos Planos Individuais	Foram criados 62 instrumentos e atingimos uma taxa de execução de Pl's de 77,75%, o que corresponde a um aumento de 16%. O objetivo foi atingido a 100%.	NA

Proporcionar aos clientes atividades que lhes permitam a concretização dos objetivos dos PDI	Participar no desenvolvimento do projeto individual dos clientes	Participação no desenvolvimento de 73 PDI Cumprimento de pelo menos 70% das atividades	Nº de PDI em que se participou % de atividades realizadas	Participação no desenvolvimento de atividades para 73 PDI. 100% de atividades realizadas (objetivo superado)	NA
Avaliar as atividades desenvolvidas, ao nível dos mecanismos colocados à disposição dos clientes	Distribuição e tratamento de instrumentos de avaliação	Pelo menos 60% dos clientes deverão estar satisfeitos com as atividades que desenvolveram	% satisfação	94% clientes satisfeitos	NA

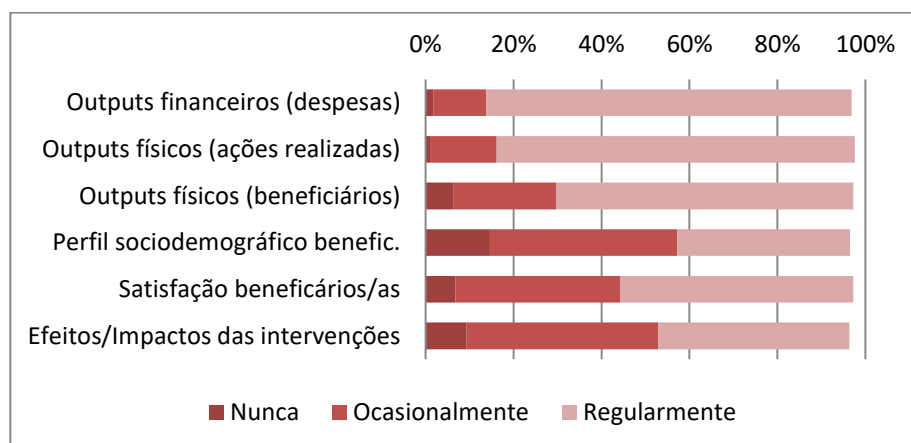
Fonte: Relatório de Atividades de IPSS - 13.

Também ao **nível da organização como um todo** prevalece, pois, uma perspectiva de **avaliação de processos**, muito assente na dinâmica interna das organizações, dos seus instrumentos e políticas de atuação, dos seus mecanismos de prestação de serviços, das suas práticas de gestão e das ligações entre estes elementos. Retomando a distinção de Osborne (1992), **vemos que o foco é colocado de na qualidade do serviço prestado e não na qualidade de vida das pessoas destinatárias das intervenções.**

A **secundarização da informação sobre impacto** nas práticas das OTS é reiterada numa interpelação do questionário em que as organizações foram desafiadas a elencar por ordem de relevância 3 tipos de preocupações que presidem aos exercícios de avaliação. “Saber o grau de satisfação das pessoas que participam nas atividades” é a 1ª opção selecionada pela maior parte das entidades (46%). Seque-se a opção “Saber quantas pessoas são atendidas, quantas pessoas participam nas atividades” (29%) e, por fim, “Saber o impacto/as mudanças produzidas na vida das pessoas”, que preside aos propósitos das avaliações de uma percentagem de entidades inferior (25%).

Noutra questão colocada no questionário sobre o tipo de informação recolhida com maior frequência as respostas reiteram a enfoque em informação que descreve os aspetos relativos à implementação das atividades, em desfavor da informação sobre os efeitos/mudanças, positivas ou negativas, de curto, médio ou longo prazo, sobre a população (vide gráfico 4).

Gráfico 4. Tipo de informação de monitorização/avaliação recolhida



Fonte: Inquérito por questionário.

Estudos sobre práticas de avaliação nas OTS conduzidos noutros países também apontam para a prevalência da informação sobre realizações/outputs relativamente à informação sobre os efeitos/impacto, ainda que os valores de adesão à avaliação do impacto encontrados em países anglo-saxónicos como o Reino Unido ou os EUA sejam muito superiores aos encontrados em Portugal e noutros países do Sul da Europa.¹¹³

Tal como análises conduzidas noutros países já haviam feito notar (e.g., Hall *et al.*, 2003), a relutância das entidades em relação à avaliação dos seus efeitos/impactos é também devedora da dificuldade e consumo de tempo e recursos associados à recolha deste tipo de informação. Na verdade, o processo de avaliação pode apresentar uma série de desafios para as organizações em termos de conceção e execução. E estes **desafios aumentam à medida que uma organização se move de uma avaliação de atividades e outputs para uma avaliação de efeitos (outcomes) ou impactos**. A avaliação dos *outputs* geralmente relaciona-se com as atividades da organização, tais como o número de ações realizadas, o número de pessoas voluntárias envolvidas, ou o número de clientes atendidos/as. As informações necessárias para fazer este tipo de avaliação são relativamente fáceis de recolher. Já a avaliação dos efeitos é frequentemente vista como mais difícil de fazer, porque implica que as organizações encontrem forma de analisar

¹¹³ Num estudo conduzido em Itália (Melloni, 2014), por exemplo, apurou-se que menos de um terço das OTS avaliavam os efeitos/impactos das suas atividades numa base regular, enquanto estudos similares conduzidos no Reino Unido (Lyon, McKay e Moro, 2013; Ógáin, Lumley e Pritchard, 2012) e nos EUA (Salamon, Geller e Mengel, 2010) apontam para valores de adesão à recolha de informação sobre benefícios diretos ou indiretos das intervenções que rondam os três quartos das organizações.

comportamentos que ocorrem depois de um serviço ser prestado ou de uma atividade ter decorrido. Isto implica definir as consequências desejadas para as pessoas que beneficiam dos serviços ou atividades (por exemplo, através da teoria da mudança), encontrar uma forma de mensurar essas consequências, e manter o contacto com as pessoas beneficiárias, as quais poderão já não estar ativamente envolvidas com a organização, a fim de monitorar os impactos (Hall *et al.*, 2003: 17).

Nas entrevistas realizadas com representantes de OTS, é clara a pouca ou nenhuma atenção dedicada a análises dos efeitos/impactos na generalidade das entidades. O excerto abaixo é paradigmático:

Nós acabamos por privilegiar sobretudo isso [os *outputs*], porque não há a tal avaliação de impacto, para além daquilo que é aqueles números de execução do projeto. Nem sempre há capacidade da avaliação de fazer uma avaliação de impacto ao nível do contexto onde nós trabalhamos. Acabamos por ficar por uma avaliação muito circunscrita à execução das metas. (Entrevista com dirigente de OTS – 20)

Mas, no caso português, aos **obstáculos associados à capacidade das organizações para fazer este tipo de análise** (em termos de recurso e competências) acrescenta-se a **perceção do desinteresse das entidades financiadoras** e em relação a este tipo de investimento por parte das entidades que financiam:

É mais fácil ver em termos de resultados. Depois os impactos... depende, depende também... gostaríamos que desse ponto de vista fosse mais ambiciosa. Mas aqui a questão é sempre do tempo, não é uma desculpa, mas é um facto. (...) E depois também... os financiadores, as propostas não têm em conta a avaliação que fizemos. Os próprios programas que fazem avaliações *ex-ante* e *ex-post*, que têm essa forma de trabalhar, também não incorporam nas novas propostas muitos dos resultados dessa avaliação... (Entrevista com coordenadora/técnica de ADL – 26)

Algumas análises (Hall *et al.*, 2003; Ellis and Gregory, 2008; Hoefler, 2009; Benjamin, 2008, 2012) dão ainda conta do aumento substancial ao longo dos últimos anos do número de organizações que medem o seu impacto, passando a dedicar uma crescente proporção do orçamento na recolha e análise deste tipo de informação (Ellis e Gregory, 2008). Mas o **reforço do investimento neste tipo de análise** em países como o Reino Unido, os EUA ou o Canadá, **surge muito associado ao crescente interesse das entidades financiadoras em avaliações que incidam sobre as mudanças produzidas pelos programas e projetos que financiam** face às avaliações que reportam sobre atividades e realizações/*outputs* dos programas ou projetos (Hall *et al.*, 2003; Ellis and Gregory, 2008;

Benjamin, 2008)¹¹⁴. Ora, **em Portugal, o “novo” interesse pelo impacto (particularmente associado à abordagem da filantropia de impacto) tem ainda uma história muito curta, não sendo ainda perceptíveis os seus efeitos nas organizações**¹¹⁵.

Encontraremos aqui algum fundamento para o facto de as **práticas que as organizações mais associam à priorização da recolha de informação sobre o impacto**, nomeadamente a abordagem SROI (ferramenta de avaliação fortemente orientada para a medição do impacto social), **tendam a ser dissociadas das pressões coercivas do financiamento**¹¹⁶, ao contrário do que acontece em outros países em que o “mantra do impacto” (Ebrahim e Ragan, 2010: 2) tem maior histórico de implementação, como é o caso do Reino Unido. Estudos neste país dão já conta da enorme relevância das pressões coercivas exercidas pelas entidades financiadoras sobre as OTS no sentido de conduzirem exercícios de medição do impacto mais rigorosos – designadamente através de abordagens “value for money”, como é o caso do SROI (Ógáin, Lumley e Pritchard, 2012, Arvidson, 2011; Harlock, 2013; Arvidson e Lyon, 2014).

Em termos gerais, a análise dos perfis de entidade associados aos diversos enfoques/conteúdos da avaliação¹¹⁷ releva a influência particularmente forte na adesão a

¹¹⁴ Num inquérito a OTS do Reino Unido liderado por Lyon, McKay e Moro (2010, 2013), mais de metade das entidades que dizem ter aumentado o seu investimento na medição dos efeitos/impactos nos últimos 5 anos, fizeram-no para atender às exigências das entidades financiadoras. Apesar do crescente interesse destas entidades na capacidade das organizações demonstrarem “a diferença que fazem” face à capacidade que têm para “executar as atividades previstas”, é também enfatizado nestes estudos o desconforto ou dissociação de muitas organizações em relação a este “novo” interesse (Arvidson e Lyon, 2014).

¹¹⁵ Esta orientação, ainda em fase embrionária aquando do trabalho de campo, passou a ter uma face mais visível a partir do novo quadro comunitário de apoio, o Portugal 2020. A realização do trabalho de campo foi feita quando ainda vigorava o QREN.

¹¹⁶ Depois de se estabelecerem correlações entre práticas de gestão e avaliação (binominais: 1 = nunca/ocasionalmente; 2 = regulamente) e a variável “priorização de informação sobre os impactos das intervenções” (binominal: 1 = Não; 2 = Sim), através do coeficiente de correlação Phi (ϕ), foram correlacionadas as práticas que apresentavam associações com significância estatística ($<0,05$) à variável “Fazemos avaliação porque os financiadores exigem”. Para todos os processos avaliativos alvitrados, o sentido das correlações foi negativo, indicando dissociação com a afirmação. Coeficientes de correlação ϕ : avaliações formais de desempenho organizacional: $\phi = - 0,150$; monitorização de objetivos e metas organizacionais: $\phi = - 0,118$; SROI: $\phi = 0,110$; modelos lógicos: $\phi = - 103$. Todas estas associações são estatisticamente significativas ($<0,05$.)

¹¹⁷ Para avaliar se as modalidades de avaliação, em termos dos seus conteúdos, variam segundo as principais características do contexto e estrutura organizacional (tipo de organização, funções principais, fonte de receitas dominante, dimensão e antiguidade) recorreu-se ao Teste de Qui-quadrado. Observou-se, desde logo, que estas características são particularmente importantes para compreender as preocupações das organizações com a aferição das mudanças produzidas na vida das pessoas. No topo da influência da

alguns tipos de informação de avaliação da antiguidade das organizações e do seu perfil de receitas, mas também do perfil da sua atividade. Encontramos uma **maior abertura das organizações de constituição mais recente, particularmente das que dependem sobretudo de donativos e com um perfil de funções misto ao investimento na recolha de informação sobre os efeitos mediatos e indiretos** da sua intervenção, ou seja, à avaliação de impacto. As entidades de constituição menos recente e autossustentadas ou dependentes de receitas públicas, tendem a privilegiar padrões de avaliação de processo, favorecendo a recolha de informação sobre realizações físicas e financeiras e/ou sobre a satisfação com os serviços/atividades desenvolvidos. O padrão de avaliação de processo é também prevacente nas **entidades com funções de expressão**, que tendem a centrar a sua análise nos *outputs* e a **dissociar-se da recolha de informação sobre o impacto** da sua intervenção. **Esta observação reforça a ideia do desafio acrescido que se coloca às organizações de advocacia ou expressão de interesses para medir os efeitos da sua ação, veiculada noutras análises** (Bryson, 2011; Eckert e Moulton, 2011; Højlund, 2014). A menor propensão destas organizações para a avaliação do impacto prende-se exatamente com carácter intangível das mudanças que geram, que torna difícil a sua aferição, e com as dificuldades acrescidas na atribuição dos resultados à sua ação, face ao débil controlo que tendem a ter sobre os efeitos (Ebrahim e Regan, 2010).

É também significativa para a análise a **relevância modesta que tem o fator “dimensão financeira” na fundamentação da adesão à avaliação de impacto**. Isto terá que ver, em parte, com a associação da avaliação de impacto à avaliação conduzida no âmbito de projetos específicos, a qual é mais integradora de entidades com diferentes níveis de capacidade financeira. Tal significa que, apesar de se tratar de um tipo de avaliação exigente, ele não é necessariamente exclusivo de grandes organizações.

adesão a este tipo de informação está a antiguidade da organização, a sua principal fonte de receitas (ambas com $p = 0,000$), a que se seguem o tipo de entidade ($p = 0,004$) e as principais funções ($p = 0,007$). Com efeito, são as entidades de constituição mais recente, dependentes de filantropia, ONG e ADL, e com funções mistas, que mais priorizam o conhecimento sobre o **impacto** da sua atividade na vida das pessoas. A **satisfação**, por seu turno, é significativamente influenciada pela antiguidade ($p = 0,003$) e principal fonte de receitas das organizações ($p = 0,009$), encabeçando sobretudo as preocupações das entidades de constituição menos recente (+ 20 anos) e sustentadas por receitas públicas ou próprias. A preocupação com os **outputs/realizações** é condicionada, de forma estatisticamente significativa, pelo perfil da atividade ($p = 0,011$) e pela antiguidade (0,013), abrangendo significativamente mais organizações com funções de expressão e sendo mais comum nas entidades de constituição mais antiga (+ 50 anos).

Outra nota importante respeita à **associação da avaliação de impacto às entidades cuja principal fonte de receitas advém da filantropia** (doações individuais, doações de fundações e doações empresariais). Esta **associação é coerente com o enquadramento da recente atenção dedicada à avaliação/medição do impacto das OTS no nosso país**. Como vimos atrás, as fundações filantrópicas e empresariais têm vindo a privilegiar a demonstração do impacto social dos projetos que financiam, e têm marcado fortemente a agenda sobre o estabelecimento de *standards* e metodologias para a análise do impacto. Um dos marcos desta presença foi o lançamento por uma fundação empresarial – a Fundação Montepio – e pela CASES de um programa de formação para ajudar as entidades da Economia Social a demonstrar o impacto social e económico que criam através de uma ferramenta de análise de custo-benefício – o SROI. Entretanto, sobretudo com base nesta mesma metodologia, têm-se multiplicado a oferta de serviços e de formação apoiados por organizações filantrópicas.

Síntese

Embora a filosofia de gestão baseada em resultados pareça permear os discursos e práticas de uma parte importante de organizações, a conceção de resultados que prevalece nas organizações vincula-se quase exclusivamente aos produtos diretos dos serviços ou das atividades das organizações (frequência das ações e volume de participantes/taxas de execução). É muito menos frequente a conceção de resultados enquanto tradução dos efeitos das intervenções, mesmo dos efeitos que é possível identificar de forma direta e imediata junto das pessoas destinatárias. O reforço do investimento por parte das OTS em análises que incidem sobre mudanças em países onde já têm expressão significativa tem surgido muito associado ao crescente interesse das entidades financiadoras. Mas, no caso português, aos obstáculos associados à capacidade das organizações para fazer este tipo de análise (em termos de recurso e competências) acrescenta-se a perceção do desinteresse das entidades financiadoras e em relação a este tipo de investimento por parte das entidades que financiam. A ausência desta perspetiva de longo prazo espelha, pois, a própria orientação dos requisitos do financiamento, ainda limitados no propósito de direcionar os processos de monitorização e avaliação das intervenções para as mudanças. Apesar de esta tendência se verificar ao nível dos

requisitos do financiamento público para a implementação de projetos, as experiências sobre a condução de avaliações de impacto têm sobretudo lugar no âmbito de projetos específicos. De resto, os exercícios formais de avaliação, mais enformados pelo campo da pesquisa em avaliação, ocorrem sobretudo a este nível – e quando há financiamento específico para o efeito. É principalmente no âmbito de projetos que se podem encontrar as análises mais sofisticadas e exigentes (em termos técnicos) do desempenho das intervenções: uma maior diversidade e riqueza de metodologias (qualitativas e quantitativas), um maior envolvimento das partes interessadas e um maior recurso a especialistas em avaliação.

4.4.3. Participantes na avaliação: o fraco envolvimento dos stakeholders

As abordagens participativas para a avaliação podem ser diferenciadas de outras abordagens de avaliação (e entre si) com base no nível e extensão do envolvimento das partes interessadas, na diversidade de grupos envolvidos e no nível de controlo de quem avalia sobre o processo (Cousins e Whitmore, 1998).

Os fundamentos comumente invocados para o recurso a abordagens participativas/colaborativas incluem justificações políticas, filosóficas e pragmáticas, sintetizadas no trabalho de Jill Chouinard (2013: 242). O fundamento político – que se pode associar ao paradigma transformativo – encontra-se alicerçado nas preocupações com a justiça social e baseia-se num sentido moral e normativo de obrigação que envolve os ideais de inclusão democrática, *empowerment* e emancipação (Cousins e Whitmore, 1998). A justificação filosófica apoia-se no construtivismo social – paradigma construtivista – e refere-se à centralidade do contexto e à inclusão de múltiplos participantes na produção do conhecimento. Por fim, a justificação pragmática¹¹⁸ – que se pode associar ao paradigma pragmático –, com uma orientação prática, baseia-se na crença de que o aumento da participação conduzirá a resultados que melhor suportarão a tomada de decisão organizacional (Patton, 1998) e a aprendizagem organizacional (Torres e Preskill, 2001).

¹¹⁸ Os efeitos da participação dos *stakeholders* nos usos da avaliação será explorado no capítulo seguinte.

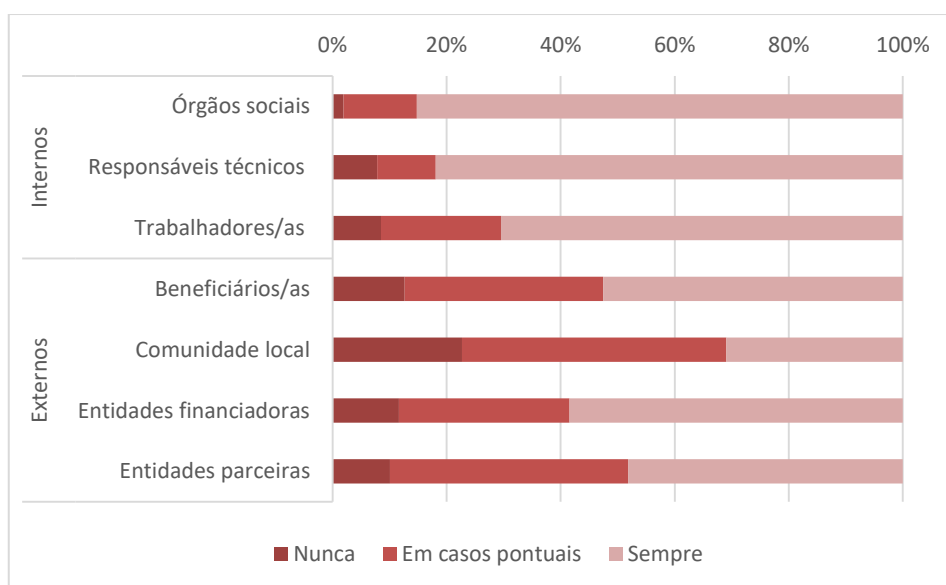
Essas três justificações, embora não sejam conceitualmente distintas, transmitem uma sensação de múltiplas motivações para adotar uma abordagem participativa de avaliação.

Embora seja generalizado o reconhecimento da relevância do envolvimento das diversas partes interessadas na avaliação, **ele aparece situar-se mais no plano discursivo dos desejos e aspirações do que em práticas abrangentes, estruturadas e aprofundadas de participação**, à semelhança do que concluíram Ruela e Albuquerque (2016) no seu estudo. Quando no inquérito por questionário as OTS são interrogadas sobre quem geralmente participa nos processos de avaliação, os *stakeholders* internos são predominantes: membros da direção (92,6%) e/ou outros órgãos de governação (52,8%), em primeiro lugar, e responsáveis técnicos (74,5%) e/ou outros/as trabalhadores/as (63,5%) em segundo, o que denota alguma **concentração dos processos de avaliação na componente técnica, mais restrita aos profissionais e chefias de topo**. Só depois surgem *stakeholders* externos: utilizadores ou beneficiários (58,7%), entidades financiadoras (37,4%) e, finalmente, comunidade local (31,8%), realçando o carácter fechado dos processos de avaliação.

Além de serem os principais participantes nos processos de avaliação, os **órgãos de governação** e as **pessoas que trabalham** nas organizações (responsáveis técnicas e outras) são também os principais destinatários das avaliações produzidas. Ou seja, **a informação de avaliação não só é participada mas também comunicada, sobretudo, a nível interno. A nível externo, as entidades financiadoras/tutelares são as destinatárias privilegiadas**. Só depois surgem os públicos-alvo das organizações, acompanhados de perto das entidades parceiras e, a alguma distância, a comunidade mais alargada. E aqui encontramos evidência acrescida do **carácter eminentemente ascendente da responsabilização**, dada a descoincidência entre a relevância da participação de entidades financiadoras e de pessoas beneficiárias (e até da comunidade local) e a proeminência destes grupos de *stakeholders* ao nível da comunicação dos resultados de avaliação. Com efeito, enquanto a presença das entidades financiadoras em termos de comunicação dos resultados da avaliação é assimetricamente superior à participação nos exercícios de avaliação, no caso de pessoas beneficiárias e da comunidade local acontece o inverso. Tal significa que a devolução dos resultados da avaliação às populações que servem e que são

de algum modo envolvidas no desenho e/ou implementação não está a ser feita de forma plena. **A capacidade de exigir prestação de contas não é, no entanto, uniforme, sendo a mais forte a dos financiadores e a mais fraca a dos beneficiários, o que sinaliza assimetrias de poder e o enviesamento dos processos de avaliação em favor dos atores dominantes.**

Gráfico 5. Comunicação dos resultados de avaliação



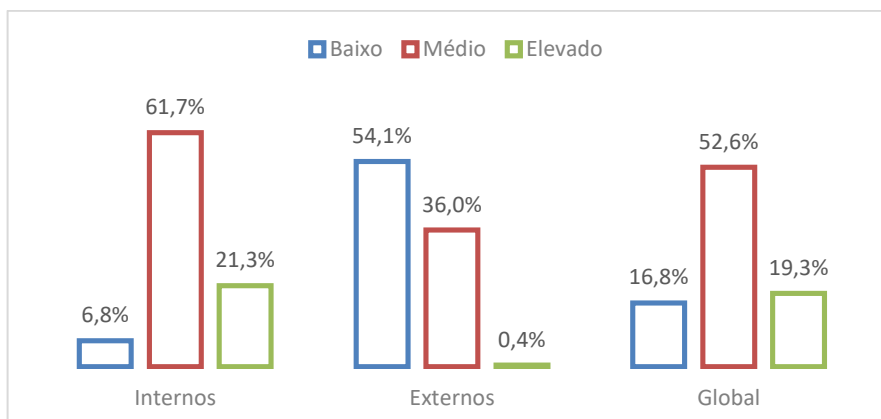
Fonte: Inquérito por questionário.

A organização da informação sobre a participação de *stakeholders* segundo níveis de intensidade¹¹⁹ permite-nos alargar a perspetiva da análise da participação. Vemos que **os níveis globais de participação (interna e externa) são medianos, mas sobretudo devedores da participação interna. A participação de pessoas exteriores apresenta níveis**

¹¹⁹ Níveis de participação definidos a partir do número de grupos de *stakeholders* envolvidos pelas organizações na conceção/implementação dos exercícios de avaliação. Para a participação interna e externa foi usada a seguinte escala: 0 a 1 grupo de *stakeholders* (nível baixo); 2 a 3 (nível médio); e 4 ou mais (nível elevado). Para a participação global (interna e externa) foi usado uma escala diferente: 0 a 2 grupos de *stakeholders* (nível baixo); 3 a 5 (nível médio); e 6 ou mais (nível elevado).

tendencialmente baixos, evidenciando a dificuldade de abertura dos processos de avaliação ao conjunto mais alargado de atores implicados na intervenção das OTS.

Gráfico 6. Níveis de participação de *stakeholders* na avaliação



Fonte: Inquérito por questionário.

A partir da análise do inquérito por questionário tentou perceber-se qual a abordagem metodológica mais propícia ao maior envolvimento das partes interessadas e quais os dispositivos metodológicos privilegiados pelas organizações envolver os diversos grupos de *stakeholders*. Observa-se, desde logo, uma correlação positiva estatisticamente significativa entre a diversificação dos métodos de recolha de dados para avaliação, particularmente de informação primária, e o nível de envolvimento de *stakeholders* na avaliação, **exaltando as vantagens das abordagens multi-método para o envolvimento alargado das partes interessadas**¹²⁰. Por outro lado, à exceção da recolha de dados preexistentes (administrativos e documentais), todas as restantes técnicas de recolha de informação se correlacionam positivamente com a diversificação do envolvimento de *stakeholders* (internos e externos) na condução de exercícios avaliativos (estudos de caso, entrevistas individuais, *focus groups*, inquéritos, grupos de trabalho de avaliação e observação estruturada). **Uma metodologia de avaliação assente na recolha de dados**

¹²⁰ Para avaliar se a abrangência da participação de *stakeholders* na avaliação (número de grupos de *stakeholders* internos/externos envolvidos) se correlaciona com a diversidade de métodos de recolha de informação (número de métodos mobilizados) recorreu-se ao coeficiente de correlação de Pearson (variáveis medidas numa escala quantitativa). Para avaliar as correlações entre a participação de cada um dos grupos de *stakeholders* na avaliação e a adesão a cada um dos métodos de recolha de informação recorreu-se ao coeficiente de correlação *Phi* (variáveis medidas numa escala nominal). Para avaliar as correlações entre a abrangência da participação de *stakeholders* e a adesão a cada um dos métodos de recolha de informação recorreu-se ao coeficiente *Eta* (variáveis medidas numa escala nominal *versus* variáveis medidas numa escala quantitativa) (vide tabela 5, anexo 5).

administrativos é, pois, como seria de esperar, a menos rica em termos de dinâmicas de participação. Dentre as diversas técnicas de recolha de informação primária, o inquérito parece ser a forma privilegiada de acesso aos contributos de atores externos para a avaliação, enquanto os grupos de trabalho de avaliação (que em muitos casos não serão grupos de trabalho específicos, mas reuniões de trabalho onde a avaliação é um dos tópicos abordados) são os momentos eleitos para envolver um maior número de grupos de *stakeholders* internos nos processos avaliativos. Nas entrevistas com representantes de OTS é clara a centralidade do inquérito por questionário, nomeadamente do inquérito de satisfação, enquanto veículo de envolvimento de *stakeholders* externos na avaliação – especialmente de participantes em eventos ou atividades formativas e de utentes/utilizadores/as dos serviços e, menos frequentemente, de parceiros e da comunidade local. A mobilização de metodologias qualitativas para o envolvimento de atores externos, nomeadamente de entrevista individuais ou de grupo, é muito menos frequente, ainda que geralmente consideradas mais profícuas nos seus resultados:

Sem carácter regular, há outras ferramentas... às vezes faço entrevistas que gravo, que duram no mínimo uma hora, em que faço 10 perguntas sobre todas as questões. (...) Faço uma amostra aleatória e converso com eles durante um bom pedaço. E é aí que tenho informação absolutamente deliciosa. (...) É a melhor metodologia. (Entrevista com coordenação técnica de IPSS – 5)

De resto, em termos globais, as análises de correlação entre as práticas de controlo, gestão e avaliação (e respetiva intensidade/abrangência) e níveis de participação de *stakeholders* (e com cada um dos grupos de atores) nos exercícios avaliativos, são reveladoras da forte associação entre níveis de participação, particularmente os externos, e níveis de intensidade/abrangência de práticas de gestão, como se vê na tabela abaixo¹²¹.

¹²¹ Para avaliar se a abrangência da participação de *stakeholders* na avaliação (número de grupos de *stakeholders* internos/externos envolvidos) se correlaciona com o nível de adesão a cada uma das práticas de gestão) recorreu-se ao coeficiente de correlação de Spearman (variáveis medidas numa escala quantitativa e ordinal, respetivamente). Para avaliar se a abrangência da participação de *stakeholders* na avaliação (número de grupos de *stakeholders* internos/externos envolvidos) se correlaciona intensidade das práticas de gestão (n.º de práticas) recorreu-se ao coeficiente de correlação de Pearson (variáveis medidas numa escala quantitativa).

Tabela 6. Correlações entre práticas de controlo, gestão e avaliação e níveis de participação global, interna e externa na avaliação

	Nível de participação	Nível de participação interna	Nível de participação externa
Relatórios anuais de atividades	0,07	0,057	0,039
Relatórios atividades de projetos	0,235*	0,127*	0,247*
Relatórios execução financeira de projetos	0,259*	0,169*	0,266*
Avaliações formais de projetos	0,300*	0,186*	0,323*
Processos de certificação	0,270*	0,207*	0,289*
Auditorias financeiras externas	0,278*	0,194*	0,285*
Auditorias técnicas externas	0,335*	0,239*	0,323*
Avaliação desempenho pessoal	0,318*	0,284*	0,273*
Monitorização de objetivos e metas organizacionais	0,231*	0,155*	0,276*
Planeamento estratégico	0,276*	0,218*	0,288*
Balanced scorecard	0,238*	0,18*	0,254*
SROI – Social Return on Investment	0,183*	0,153*	0,171*
Avaliações desempenho organizacional	0,318*	0,240*	0,292*
Modelos lógicos	0,332*	0,224*	0,371*
N.º médio de práticas	0,395*	0,264*	0,423*

Fonte: Inquérito por questionário.

Nota: * A correlação é significativa no nível 0,05 (bilateral).

Em termos globais, as práticas **mais mobilizadoras da participação** são as vinculadas a **auditorias técnicas**/visitas de acompanhamento da iniciativa de entidades financiadoras/tutelares e **exercícios de avaliação formal** à escala individual, organizacional, e dos projetos (especialmente os estruturados a partir de modelos lógicos). Não deixa de ser **reveladora da dificuldade das organizações em promover a participação alargada nos processos de gestão e avaliação** o facto de entre os processos com maior envolvimento das partes interessadas se encontrem os dinamizados por iniciativa de atores externos às OTS. O reforço dos níveis de participação externa, em particular, surge especialmente associado a exercícios estruturados de avaliação à escala dos projetos e as auditorias técnicas externas, enquanto o dos níveis de participação interna surgem especialmente associados à avaliação de desempenho individual.

Esta análise revela já o contributo da condução de exercícios de avaliação estruturados, sobretudo à escala dos projetos, para a abertura da avaliação à participação de atores externos às organizações¹²². De resto, a discussão sobre metodologias participativas de avaliação tem estado particularmente associada ao contexto do desenvolvimento de programas e projetos. Aliás, muita do debate teórico em torno de abordagens participativas/colaborativas à avaliação tem emergido da literatura sobre avaliação de programas (Green, 1988; Cousins e Earl, 1993; Cousins e Whitmore, 1998; Fitzpatrick, Sanders e Worthen, 2011; Cousins, 2013). No sentido de corroborar esta análise, será também relevante perceber o efeito estatístico do envolvimento em projetos nos níveis de participação dos diversos *stakeholders* na avaliação.

A análise das correlações entre níveis de participação de *stakeholders* na avaliação e o envolvimento em projetos são reveladoras de associações significativas (vide tabela 6, anexo 5). O **envolvimento em projetos é globalmente potenciador da abertura à participação de um conjunto mais alargado de *stakeholders*, sobretudo externos às organizações**. Apesar da associação mais forte do envolvimento de projetos a níveis de participação externa, as associações individuais mais robustas (com cada um dos grupos de *stakeholders*) registam-se com grupos de *stakeholders* internos: responsáveis técnicos/as e outros/as trabalhadores/as. A **nível interno**, o envolvimento em projetos parece deslocar a **participação dos órgãos de governação estatutários para as chefias e equipas técnicas**, já que ele induz uma retração significativa do envolvimento daqueles órgãos. A nível externo, a participação em projetos é sobretudo potenciadora do maior envolvimento de beneficiários/as e entidades financiadoras, demonstrando uma reduzida capacidade de mobilização da comunidade local mais alargada.

A análise das entrevistas, por seu turno, denuncia, grosso modo, a dificuldade das organizações em promover uma **avaliação participativa – que ultrapassasse o nível superficial da consulta ao público-alvo** – desenvolvendo metodologias para que todos os *stakeholders* tenham um papel ativo nas diversas fases do processo avaliativo (desde o

¹²² Esta aceção é reforçada pelos resultados da desagregação da análise dos níveis de adesão por grupos de *stakeholders*, na qual se destaca a importância da participação de pessoas beneficiárias e, sobretudo da comunidade local, em processos e avaliação formal.

desenho até à interpretação dos resultados da avaliação). **São raros os casos em que as iniciativas de avaliação se abrem à negociação entre as diferentes partes interessadas sobre o que vai ser monitorizado e avaliado, sobre as condições em que os dados são recolhidos e analisados, o realmente significam, ou sobre a forma como os resultados serão partilhados e usados.** Registam-se, no entanto, alguns bons exemplos de organizações que constroem soluções criativas e eficazes para incluírem na avaliação pessoas beneficiárias e outros atores externos, institucionais e não institucionais, e que assentam sobretudo em técnicas qualitativas, como *focus groups*, balanço de competências, ou técnicas audiovisuais, como o photovoice. O testemunho que se segue é exemplar da procura de soluções para o envolvimento dos diversos agentes implicados na intervenção, buscando adaptar a avaliação a um ambiente externo mais amplo e a um conjunto de condições e atores locais:

O nosso modelo de intervenção é sempre em conjunto com organizações locais, enquanto organização a trabalhar no continente africano. (...) Depois, obviamente, o nosso mapa de necessidades locais... são trabalhadas perguntando às pessoas dos locais o que é que elas pensam que são os melhores caminhos para a transformação social. (...) [Os avaliadores externos] têm sempre que nos apresentar uma metodologia que nos permite ouvir os intervenientes. (...) Queremos ouvir os beneficiários, as Estado, as organizações que, não tendo estado envolvidas no projeto, estão próximas das pessoas com quem trabalhamos. (...) Nós procuramos que eles apresentem metodologias que façam análises qualitativas. (...) Eu costumo dizer que quando faço entrevistas ou FG em comunidades, o melhor de tudo acontece quando eu vou embora. E, portanto, este empoderamento, de serem as pessoas a ser o veículo de filmar e de olharem para o que entendem que é o mais significativo das mudanças que estão a acontecer com o projeto. (Entrevista com dirigente de ONG – 6)

Boa parte destas **experiências participativas** são conduzidas no âmbito de **projetos** ou passaram a ser integradas na rotina das organizações depois de serem experimentadas em projetos, como é o caso que se traz à colação:

Tudo o que os projetos fazem é um valor acrescentado que se pretende que a organização incorpore a sua estratégia. E quando se falou na escuta e envolvimento de *stakeholders*, foi um projeto que teve oportunidade de fazer essa experimentação, essa incorporação na organização que, no fundo, é uma metodologia que existe a nível internacional, que é a AA1000 [ferramenta de gestão de *stakeholders*]. Nós, nessa altura, ouvimos os 71 *stakeholders*, e os meios que nós utilizámos dependeram em função... Por exemplo, se era um *stakeholder* crítico para a organização, optamos por fazer por entrevista individual: por exemplo, o presidente da Segurança Social ou do IEFPP, fomos lá. Grupos como as famílias, clientes, universidades, como não poderia ser só 1 ou 2 ou 3, optamos por ser um *focus group*. E depois também havia os questionários [financiamento Leonardo da Vince]. (Entrevista com dirigente de IPSS – 8)

Enquanto no caso dos projetos POPH, não é reconhecido contributo relevante em termos de integração de dinâmicas participativas nas organizações, os projetos conduzidos no âmbito do **PIC EQUAL** ressaltam no discurso de especialistas em avaliação e representantes de organizações de base entrevistados/as no estudo como emblemáticos das **abordagens participativas à avaliação**. De facto, um princípio-base da avaliação **EQUAL** (em coerência com a filosofia do programa) era o da participação de todos os agentes envolvidos na intervenção, procurando constituir-se um fator de *empowerment* para as pessoas participantes, ainda *que* esta recomendação do programa tenha sido debilmente acolhida por um grande número de projetos (Figueirôa e Pimenta, 2005). Esta experiência é trazida em contexto de entrevista para ilustrar práticas de participação profícuas conduzidas nas organizações, muito embora na maioria dos casos não se tenham mantido, pelo menos com a abrangência, regularidade e profundidade com que foram iniciadas.

Em todo o caso, apesar de os casos de envolvimento mais aprofundado e alargado de *stakeholders* na avaliação ocorram no âmbito de projetos (sobretudo de cooperação internacional e desenvolvimento local e de capacitação), as **práticas mais regulares e institucionalizadas de envolvimento de stakeholders narradas em entrevista decorrem dos processos de certificação/acreditação** para o domínio da formação ou na área da qualidade. Estes exercícios tendem a confinar-se a abordagens quantitativas de recolha e análise de informação sobre necessidades, expectativas e satisfação das partes interessadas com atividades ou serviços prestados. Embora em boa parte dos casos estes exercícios pareçam ser efetivamente capitalizados enquanto oportunidades de envolvimento e de melhoria, é também frequente serem encarados como **processos meramente formais e “automatizados”**, sem procura de soluções para um envolvimento mais aprofundado e alargado e sem integração efetiva dos seus resultados (muitas vezes não são devolvidos e sequer são objeto de tratamento e análise horizontal que permita tirar conclusões globais).

As abordagens participativas/colaborativas, fundamentadas nas preocupações, interesses e problemas da comunidade e portanto potenciadoras de um melhor alinhamento das OTS com as suas expectativas são, pois, pouco frequentes nas organizações. Prevalece uma perspetiva de avaliação que minimiza o envolvimento das partes interessadas, reduz a complexidade do contexto, e se afasta da visão da avaliação

como prática concetual destinada a gerar diálogo e facilitar a aprendizagem. A debilidade e superficialidade dos processos participativos que envolvem atores externos às organizações, nomeadamente os públicos-alvo das organizações encontra-se espelhado no centramento dos processos de monitorização e avaliação nas realizações/*outputs* que, ao contrário dos efeitos/mudanças, focam nas atividades em si e não em quem delas beneficia. Compõe também o retrato do tipo de abordagem à avaliação dominante nas organizações estudadas – tecnocrática/positivista – na qual a avaliação procura afirmar-se como um processo apolítico, imparcial, neutro, focado no controlo, centrado no cumprimento de objetivos de eficácia e eficiência, e que tem como finalidade construir uma visão objetiva dos acontecimentos (Lewis, 2001; Sharman, 2007; Ebrahim, 2009; Chouinard, 2013).

4.5. Padrões de utilização dos resultados das avaliações

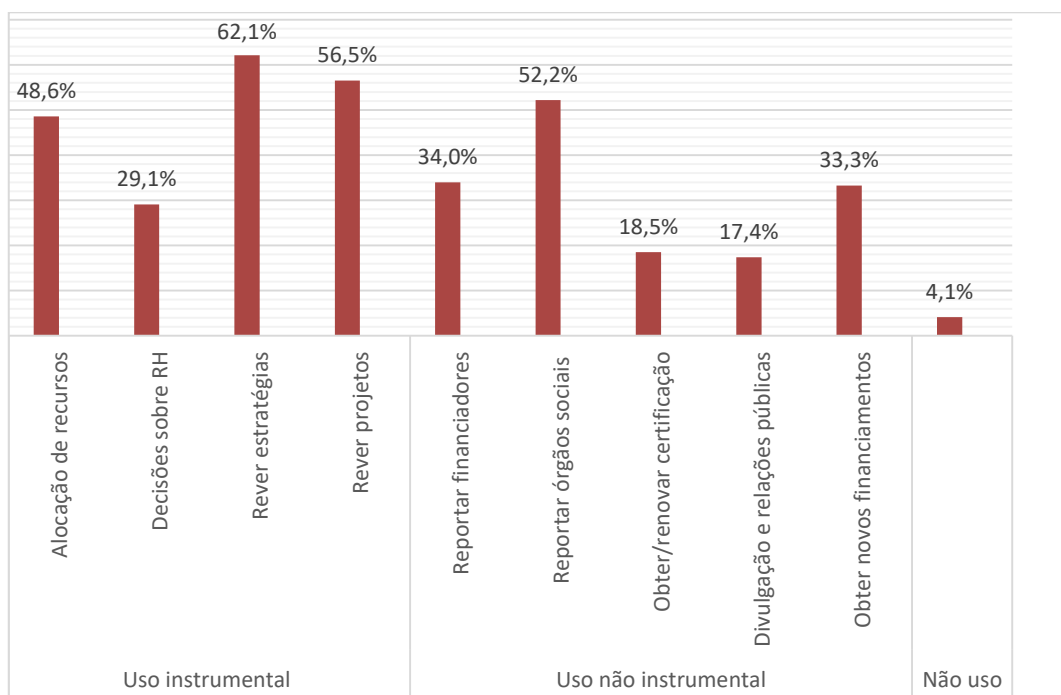
A compreensão do modo como as organizações utilizam os resultados dos exercícios de monitorização e avaliação que conduzem, assim como os diferentes tipos de uso se relacionam com o tipo de práticas adotadas e de conceções sobre avaliação, informa-nos sobre algumas das lógicas subjacentes à avaliação nas OTS.

No gráfico abaixo encontram-se sistematizadas as respostas das organizações inquiridas aos usos que fazem dos exercícios avaliativos que conduzem. Globalmente, **existe um equilíbrio entre o enfoque** nos usos que remetem para a ação direta (tomada de decisão) – **usos instrumentais** – e **os usos não instrumentais** (concetuais e simbólicos):¹²³ 9 em cada 10 organizações (91,3%) identifica pelo menos um tipo de uso instrumental, e são também a 9 em cada 10 (86,8%) que integram usos concetuais e/ou simbólicos, os quais, **não remetendo diretamente para a ação material**, podem afetar o entendimento dos *stakeholders* sobre a intervenção das organizações e potenciar a atração

¹²³ Retomo nesta análise a tipologia de usos da avaliação proposta por Alkin e Taut (2003), apresentada no primeiro capítulo do trabalho. Com efeito, o **uso instrumental** acontece quando a avaliação é usada para produzir efeito na ação direta, i.e., quando a informação é usada para modificar o objeto da avaliação; o **uso concetual** acontece quando os resultados da avaliação não induzem nenhuma ação direta, mas suscitam alterações em entendimentos concetuais específicos, decorrentes de mudanças na forma como quem utiliza a avaliação pensa sobre aspetos específicos da intervenção; e o **uso simbólico** remete para situações em que a simples existência da avaliação, e não qualquer aspeto dos seus resultados, é usada para persuadir ou convencer.

recursos, ou reforçar a reputação e demonstrar legitimidade, reivindicando uma base racional para a ação (ou inação), ou justificando posições pré-existentes. Considerados de forma desagregada, assume maior abrangência o papel instrumental da avaliação no ajustamento de estratégias (62,1%) e de projetos (56,5%). No conjunto dos usos não instrumentais a adesão de mais de metade das organizações apenas se verifica no reporte aos órgãos de governação (52,2%).

Gráfico 7. Uso dos resultados da avaliação (N=556)



Fonte: Inquérito por questionário.

Nas entrevistas, a alusão a usos instrumentais dos exercícios de avaliação é também generalizada. A aplicação instrumental dos resultados da avaliação é traduzida no suporte à tomada de decisão das organizações sobre a alocação de recursos (determinar as prioridades orçamentais), orientação e ajustamento das atividades (planeamento, revisão e melhoria das intervenções), ou sobre o quadro de pessoal (reorientar, sancionar ou premiar desempenhos).

O objetivo é sempre melhorar aquilo que possa necessitar de algum tipo de ação. Quando detetamos que há questões que podem estar a ser positivas, tentamos implementar as ações de melhoria. (Entrevista com técnica de OTS – asas)
 Fica na parte organizativa da escola, e nós servimo-nos dela para alertar os professores: “cuidado, vamos ter que fazer o máximo”. E aí temos muita ação junto dos professores. (Entrevista com dirigente de escola profissional – 10)

Lidos grosseiramente, estes resultados são coerentes com os pressupostos de racionalidade da avaliação presentes em muita da teoria sobre avaliação, inerentemente ligados à categoria de uso instrumental – à ideia da tomada de decisões racionais para adaptar a ação. Não obstante, **os usos instrumentais emergem na maioria dos casos associados a outras formas de utilização dos exercícios de avaliação mais consentâneas com tipos de uso concetual e/ou simbólico**. No inquérito, só 1 em cada 10 organizações (9,8%) conduz as avaliações apenas com propósitos de gestão interna (conforme preconiza a teoria da escolha racional), aludindo unicamente a usos instrumentais. São mais comuns usos exclusivamente não instrumentais – situação que abrange 2 em cada 10 organização (15,2%) –, ainda que a norma seja a **coexistência das duas lógicas, apontando para pressupostos de outras teorias explicativas da condução de exercícios de avaliação, como a teoria institucional, da dependência de recursos e da *stewardship***.

É comum, por exemplo, associar o emprego dos exercícios de monitorização do desempenho conduzidos no contexto de **sistemas de gestão da qualidade** em ações de **melhoria (uso instrumental) com o objetivo de demonstração da consistência do sistema para obtenção da certificação da qualidade** ou para marketing ou promoção externa (uso simbólico).

Demonstramos [à entidade certificadora] a consistência do nosso sistema de registo da atividade com vista a dotar o nível de decisão com o retorno de informação que levamos para a gestão. (...) Faz-se um grupo de melhoria em função das recomendações que elas indicam. (Entrevista com dirigente/técnico de ONG – 23)

A associação do uso instrumental de suporte à tomada de decisão à procura de legitimidade parece ser também característica das entidades que desenvolvem análises de impacto (é o caso de duas organizações entrevistadas), designadamente através da metodologia SROI. Nestes casos, além da utilidade para a melhoria dos processos internos, a avaliação é usada para visibilizar o impacto criado e/ou demonstrar que a organização está disposta a avaliar o seu trabalho, fomentando a sua reputação perante as partes interessadas. **O uso simbólico da avaliação aparece também associado ao desenvolvimento de projetos**, seja para demonstrar às entidades financiadoras a validade das iniciativas financiadas, seja para fundamentar/justificar candidaturas futuras.

4.5.1. Combinando métodos e usos da avaliação: implicações diferenciadas na cultura avaliativa

Uma vez que, como vimos anteriormente, a adesão a determinados tipos de práticas e métodos de avaliação é largamente influenciada pelo ambiente em que atuam as organizações, será razoável esperar que diferentes tipos de práticas se associem também a diferentes utilizações da avaliação. No intuito de estabelecer relações entre os tipos de práticas de avaliação adotadas (nomeadamente em termos de escala, agentes, métodos e conteúdos) e os usos que tipicamente se lhes associam, começo por estabelecer correlações entre a frequência das diferentes atividades de gestão e avaliação listadas no inquérito e os diversos tipos de usos da avaliação (vide tabela 7)¹²⁴. Enumeram-se de seguida as leituras gerais suscitadas pela sua análise.

- **As práticas de reporte descritivas associadas aos projetos** – particularmente adstritas ao tipo de adoção coercivo, como vimos atrás – **encontram-se entre as que menos se correlacionam positivamente com usos instrumentais**. Escapam a esta tendência os exercícios de avaliação estruturados à escala dos projetos (avaliações formais, modelos lógicos e SROI). As avaliações formais de projetos, concretamente, estão entre as poucas práticas que se associam de forma estatisticamente significativa a ambos os tipos de usos de uso. **Ou seja, no contexto dos projetos, apenas os exercícios de avaliação estruturados transcendem a escala dos projetos, potenciando apropriações de nível organizacional** (além de induzirem a utilização prática e direta dos resultados).

¹²⁴ Variáveis integradas na categoria “usos instrumentais”: Alocação recursos; Decisões sobre RH; Rever estratégias; Rever estratégias e Rever projetos. Variáveis integradas na categoria “usos não instrumentais”: Reportar financiadores; Reportar órgãos sociais; Obter certificação; Relações públicas e divulgação; e Obter novos financiamentos.

Tabela 7. Frequência das atividades de controlo, gestão e avaliação segundo os tipos de usos da avaliação

Frequência das atividades de gestão e avaliação (1 – Nunca, 2 – Ocasionalmente, 3 – Regularmente)	Usos instrumentais			Usos não instrumentais		
	Não	Sim	Correlações Spearman	Não	Sim	Correlações Spearman
	Média	Média		Média	Média	
Relatórios anuais de atividades	2.9	2.9	,022	2.9	2.9	,021
Relatórios atividades de projetos	2.3	2.5	,085	2.3	2.5	,143**
Rel. execução financeira de projetos	2.0	2.5	,069	2.2	2.5	,171**
Avaliações formais de projetos	1.7	2.3	,167**	2.0	2.3	,106*
Processos de certificação	1.6	1.8	,073	1.6	1.8	,120**
Auditorias financeiras externas	1.4	1.9	,116**	1.8	1.9	,109*
Auditorias técnicas externas	1.7	2.1	,137**	1.9	2.1	,091*
Avaliação desempenho do pessoal	1.4	2.1	,229**	2.0	2.1	,015
Monit. objetivos e metas organiz.	2.0	2.5	,209**	2.3	2.5	,049
Planeamento estratégico	1.6	2.0	,187**	2.0	2.0	-,021
Balanced scorecard	2.9	2.9	,141**	1.3	1.4	-,005
SROI – Social Return on Investment	2.3	2.5	,177**	1.3	1.5	-,021
Aval. formais desempenho organiz.	2.0	2.5	,253**	1.8	1.9	-,027
Modelos lógicos	1.7	2.3	,196**	1.7	1.8	,005

Fonte: Inquérito por questionário.

Notas: Variáveis dependentes usadas para as correlações de Spearman: Intensidade de usos instrumentais (escala de 0 – indicando que organização não adere a nenhum uso instrumental listado – a 4 – indicando que organização adere a todos os usos instrumentais listados) e intensidade de usos não instrumentais (escala de 0 – indicando que organização não adere a nenhum uso não instrumental listado – a 5 – indicando que organização adere a todos os usos não instrumentais listados).** A correlação é significativa no nível 0,01 (bilateral). * A correlação é significativa no nível 0,05 (bilateral).

- Também os **processos de gestão da qualidade e auditoria externa** se encontram fortemente associados a **usos não instrumentais** da avaliação:
 - a. Os primeiros concentram a sua vantagem na obtenção de **selo de qualidade** ou acreditação/certificação (ainda que também sejam capitalizados para a tomada de decisão sobre alocação de recursos);
 - b. Os segundos acumulam a associação forte à **prestação de contas ascendente** com correlações estatisticamente significativas com usos instrumentais, nomeadamente o apoio à tomada de decisão sobre alocação de recursos e sobre recursos humanos (vide tabela 7, anexo 5).
- As associações mais fortes aos **usos instrumentais** encontram-se sobretudo vinculadas a **exercícios estruturados de medição do desempenho e de avaliação à escala organizacional**. Isto significa que estes processos, conotados com a gestão do desempenho, encontra-se **particularmente conectados com a tomada**

de decisão, permitindo que os ensinamentos da avaliação sejam efetivamente transferida para o ambiente organizacional.

Os resultados das correlações entre o tipo de informação de avaliação priorizada pelas organizações e os usos da avaliação acrescentam a estas leituras (tabela 8, anexo 5):

- Enquanto a priorização de **informação quantitativa** se associa à **prestação de contas ascendente**, a priorização da **informação qualitativa**, estando associada à **adoção voluntária da avaliação**, como vimos atrás, dissocia-se da utilidade não instrumental da avaliação.
- Enquanto a priorização do enfoque nas **realizações** se encontra globalmente vinculada a **usos não instrumentais**, particularmente ao reporte aos órgão de governação, a priorização do enfoque nos **efeitos** das intervenções encontra-se sobretudo associada a **usos instrumentais**, nomeadamente à revisão dos projetos.
- A priorização da análise da **satisfação** encontra-se especialmente vinculada à obtenção de **certificação**, mas, ao mesmo tempo, particularmente desvinculada de outro usos não instrumentais, como seja a atração de recursos.

A análise vai ao encontro dos pressupostos das análises conduzidas noutros estudos (Carman, 2011; Højlund, 2014) e permite-nos, desde já, contrariar a hipótese aventada por Eckerd e Moulton (2011) segundo a qual a adoção de determinados métodos de avaliação não está particularmente vinculada aos usos da avaliação. O autor e a autora propunham que as organizações adotariam determinadas práticas de avaliação simplesmente porque elas se encontravam institucionalizadas, mas que depois utilizariam os métodos de avaliação independentemente de quais fossem, para corresponder às suas necessidades estratégicas com base na heterogeneidade de papéis da organização. Vemos que, tal como os resultados do estudo daquele/as autores/as acabaram por demonstrar, **existe uma relação significativa entre a adoção de determinado tipo de práticas de avaliação e tipos de usos particulares** que se fazem daqueles exercícios.

A análise das **relações entre métodos e usos da avaliação e as percepções** sobre estes exercícios constitui, por outro lado, um contributo relevante para aferir sobre o tipo de práticas mais favorável ao desenvolvimento de uma **cultura de avaliação** assente numa atitude individual e coletiva de compreensão, aceitação, promoção e valorização da

avaliação enquanto processo contínuo de contributo para a mudança no sentido de melhoria da eficácia das organizações.

Na tabela 8, começo por estabelecer correlações estatísticas entre os vários conteúdos e enfoques da avaliação e um conjunto de afirmações que refletem os posicionamentos das entidades sobre diversos aspetos relativos ao enquadramento e impacto da avaliação nas respetivas organizações.¹²⁵ Globalmente, observa-se o seguinte:

- A prevalência do “**paradigma quantitativo**” encontra-se vinculada a **percepções mais negativas** sobre a integração e implicações da prática avaliativa, nomeadamente mediante a visão da avaliação enquanto desvio do foco essencial da intervenção.
- O **centramento na dimensão numérica das realizações** encontra-se fortemente associado à **percepção de inutilidade instrumental da avaliação**. Encontra-se também associado à ausência de enquadramento da avaliação em **sistemas de gestão mais amplos**, como o planeamento estratégico.
- O **centramento das análises nas mudanças** tende a **dissociar** as organizações de **opiniões negativas** sobre a avaliação e a **ressaltar a sua utilidade estratégica**.
- O enfoque na **satisfação** ressalta **vantagens de utilidade** não evidenciadas no exercício estatístico anterior. A associação da recolha de informação sobre satisfação à qualidade dos serviços oferecidos é devedora da centralidade da monitorização da satisfação de clientes/utilizadores e outras partes interessadas nos sistemas de gestão da qualidade, mobilizadores das práticas de mais de um quarto das entidades inquiridas, como vimos em ponto anterior.

¹²⁵ São apenas expostas na tabela as afirmações em que se registam correlações estatisticamente significativas. Ficam excluídas as seguintes: “A avaliação é parte integrante das nossas práticas de gestão”; “Os requisitos de avaliação são sobretudo obstáculos que os financiadores nos impõem”; e “O tempo e o dinheiro que gastamos em avaliação não compensam”.

Tabela 8. Correlações entre tipo de informação de avaliação privilegiada e percepções sobre avaliação

	Info. quantitativa	Info. qualitativa	Priorização <i>outputs</i>	Priorização satisfação	Priorização impacto
☛ A aval. é uma componente essencial de nossos processos de planejamento estratégico	-,073	,075	-,093*	,056	,052
☛ A avaliação ajuda-nos a fazer escolhas estratégicas sobre o futuro da organização	-,060	,058	-,095*	,033	,097*
☛ A aval. ajuda-nos a melhorar a qualidade dos serviços	-,065	,063	-,109*	,149**	,009
☛ Muito do que fazemos para a avaliação é apenas simbólico	0,066	-0,065	0,078	-0,028	-0,048
☛ Ocupar tempo e recursos na aval. tira-nos disponibilidade p/ o + importante: prestar serviços	,109*	-,108*	,067	,023	-,076

Fonte: Inquérito por questionário.

Notas: ** A correlação é significativa no nível 0,01 (bilateral). * A correlação é significativa no nível 0,05 (bilateral).

Por outro lado, a observação da tabela 9 indica a existência de correlações fortes entre as percepções sobre a integração e a relevância da avaliação nas organizações com os tipos de usos (ou não uso) dos seus resultados. Em termos gerais, vemos que:

- A **combinação de usos instrumentais e não instrumentais** revela-se mais **vantajosa para a internalização das práticas de avaliação**, enquanto os **usos exclusivamente simbólicos** ou conceptuais e o **não uso** da avaliação se associam às representações mais **negativas** e à percepção de **inutilidade instrumental** destes exercícios.
- Os usos exclusivamente não instrumentais encontram-se particularmente associados à visão da avaliação enquanto **fardo imposto** pelo financiamento, na linha dos preceitos **da teoria da agência**.

Tabela 9. Correlações entre tipos de usos da avaliação e percepções sobre avaliação

	Instrum. + não instrum.	Só uso instrumental	Só uso não instrumental	Não uso
☞ A avaliação é uma componente essencial de nossos processos de planejamento estratégico	0,035	0,058	-0,069	-0,135**
☞ A avaliação é parte integrante das nossas práticas de gestão	0,114**	-0,023	-0,121**	-0,073
☞ A avaliação ajuda-nos a fazer escolhas estratégicas sobre o futuro da organização	0,169**	0,065	-0,266**	-0,213**
☞ A avaliação ajuda-nos a melhorar a qualidade dos serviços que oferecemos	0,115**	0,003	-0,107**	-0,178**
☞ Muito do que fazemos para a avaliação é apenas simbólico	-0,171**	0,041	0,152**	-0,171**
☞ Ocupar tempo e recursos na aval. tira-nos disponib. p/ o mais imp.: prestar serviços	-0,136**	0,044	0,12	0,188**
☞ Os requisitos de avaliação são sobretudo obstáculos que os financiadores nos impõem	-0,091**	0,006	0,159**	0,045
☞ O tempo e o dinheiro que gastamos em avaliação não compensam	-0,077	-0,053	0,114**	0,197**

Fonte: Inquérito por questionário.

Notas: ** A correlação é significativa no nível 0,01 (bilateral). * A correlação é significativa no nível 0,05 (bilateral).

Estes resultados são coerentes com as visões críticas colhidas nas entrevistas com OTS, sobre a (ir)relevância para as organizações de determinado tipo de informação privilegiada por entidades financiadoras, nomeadamente no âmbito de projetos do FSE: informação quantitativa descritiva sobre indicadores financeiros e de realizações/outputs, sem incidência na compreensão das mudanças geradas pelas intervenções financiadas.

4.5.2. A influência do envolvimento dos *stakeholders* nos usos da avaliação

Vários/as autores/as têm escrito sobre a importância e o valor do envolvimento dos vários *stakeholders* no processo de avaliação e discutido os benefícios dessa participação (por exemplo, Cousins e Earl, 1992; Guerra, 2002; Fine, Thayer e Coghlan, 2000; Preskill e Torres, 2000; Torres e Preskill, 2001; Alkin e Taut, 2003). O "uso de processo", expressão cunhada por Patton, é uma consequência imputável à participação das partes interessadas nas atividades de avaliação. Em virtude da sua proximidade da avaliação, os *stakeholders* podem evoluir de modo independente dos conhecimentos substantivos emergentes da avaliação (Cousins *et al.*, 2004). A participação potencia o

desenvolvimento de competências de pesquisa e da capacidade de autocrítica, autodeterminação e análise sistemática ao nível dos atores individuais (Cousins *et al.*, 2004: 106). O processo de avaliação pode ainda mudar as atitudes dos/as utilizadores/as sobre a importância da avaliação ou sobre o potencial papel do envolvimento de múltiplos atores na tomada de decisões. O processo de avaliação participativa aumenta também a probabilidade de utilização dos resultados da avaliação, por exemplo, reforçando a disponibilidade para a utilização por parte dos *stakeholders* (Greene, 1988). O uso de processo é, nesta linha, **descrito como uma forma de aprendizagem individual, de equipa e organizacional**: quando as pessoas participam no processo de avaliação que é colaborativo e guiado pelo diálogo e pela reflexão, a aprendizagem ocorre não só a nível individual mas também de equipa e organizacional (Preskill e Torres: 2000:26).

As relações que se estabelecem entre participação e utilidade da avaliação, são ressaltadas no estudo por diversos atores entrevistados, nomeadamente ligados à gestão e avaliação de programas comunitários:

Eu conheço muitos casos em que a verdadeira causa da não percepção da utilidade dos exercícios de avaliação é porque as pessoas não são envolvidas, no desenho da avaliação. Se o promotor dos projetos está a fazer uma coisa que não sabe para o que é, é normal que depois não lhe dê uso. (...) É preciso contrariar esta ideia (...) promovendo a sua utilidade. (...) Isso vai questionar as próprias metodologias de avaliação, a própria forma como se avalia. (Entrevista com representante do observatório do QREN - 1)

No entanto, embora a importância da participação seja ressaltada no discurso das pessoas entrevistadas, como vimos no capítulo anterior, o envolvimento das partes interessadas nos exercícios de avaliação é pouco abrangente e aprofundado. O potencial de processos de avaliação pouco aprofundados e intensivos, seja em termos metodológicos, seja em termos de dinâmicas de participação, desencadearem entre agentes envolvidos aprendizagens que decorram dessa participação – usos de processo – será, naturalmente, reduzido. De facto, a análise das entrevistas com pessoal técnico e dirigente de OTS é, grosso modo, reveladora da **exiguidade das mudanças cognitivas ou comportamentais decorrentes da “exposição” à atividade de avaliação**, para além do âmbito restrito dos profissionais e chefias diretamente implicados na dinamização das atividades de monitorização e avaliação. Foram poucos os casos enumerados em que os processos de avaliação/medição do desempenho suscitaram nos/as potenciais

utilizadores/as a aquisição de novas competências, mudanças de atitude sobre a importância da avaliação ou sobre o potencial papel do envolvimento de múltiplos *stakeholders* nos processos de avaliação, ou mudanças de comportamento, mediante a reorientação das intervenções na sequência de questões suscitadas pela avaliação. Os casos em que tal acontece enquadram **processos de avaliação estruturados**, aprofundados, colaborativos e guiados pela reflexão e pelo diálogo ao longo das diversas fases. O caso que se segue, da única entidade que conta com uma pessoa dedicada a tempo inteiro exclusivamente à monitorização e avaliação, é emblemático deste tipo de aprendizagem:

Inicialmente, tomar conhecimento do que faziam foi muito fácil. As pessoas gostam de demonstrar. À medida que eu ia aprofundando a reflexão (...) é que foram havendo diferentes posturas, influenciados por diferentes fatores. Desde fatores aos níveis das repostas sociais, ao tempo dos profissionais, e algumas características pessoais também. (...) Em duas respostas em particular senti uma valorização muito grande porque era dar um outro sentido ao sistema de avaliação, em que não é avaliar por obrigatoriedade e por automatismo de fazer relatório para o final do ano. Mas era perceber “isto aconteceu, foi com estes recursos, o que falhou, o que poderia ser feito diferente, com que recursos”. (Entrevista com técnica de IPSS – 4)

Não obstante o aparente parco contributo dos processos avaliativos conduzidos pelas entidades para o desenvolvimento de aprendizagens resultantes da participação nas atividades de avaliação, os resultados do inquérito mostram que, **pelo menos a um nível mais operacional, o envolvimento das partes interessadas aumenta a probabilidade de os seus resultados serem utilizados**, confirmando disposições defendidas por diversos/as autores/as (Fine, Colette e Coghlan, 2000; Cousins, Goh e Elliott, 2008; Fleischer e Christie, 2009; Johnson *et al.*, 2009). Assim, em termos globais, o nível de participação – externa e interna – dos exercícios de monitorização e avaliação reforça a utilização instrumental e não instrumental da avaliação: observam-se associações positivas estatisticamente fortes entre o nível de participação e o uso da avaliação. Existem, não obstante, oscilações significativas no efeito da participação segundo o tipo de uso em causa.

Na tabela abaixo, na qual se sintetizam as correlações entre usos e níveis de participação na avaliação¹²⁶, é possível notar que o tipo de utilização instrumental¹²⁷ é o que mais se associa positivamente aos níveis de participação global e interno da avaliação, seguido de perto do uso misto (instrumental e não instrumental) que é, por sua vez, o que mais se destaca na associação à participação externa. Ou seja, enquanto a **participação interna é particularmente potenciadora dos usos instrumentais da avaliação**, a **participação externa é sobretudo potenciadora dos usos mistos e não instrumentais**.

Tabela 10. Correlações entre usos da avaliação e níveis de participação de *stakeholders* na avaliação

		Nível de participação	Nível de participação interna	Nível de participação externa
Usos instrumentais	Alocação recursos	0,204*	0,141*	0,077
	Decisões sobre RH	0,170*	0,191*	-0,099*
	Rever estratégias	0,257*	0,235*	0,200*
	Rever projetos	0,184*	0,142*	0,109*
	Total	0,297*	0,294*	0,167*
Só uso instrumental		-0,129*	-0,061	-0,101*
Usos não instrumentais	Reportar financiadores	0,231*	0,124*	0,276*
	Reportar órgãos sociais	-0,216*	-0,083	-0,151*
	Obter certificação	0,238*	0,168*	0,259*
	Relações públicas e divulgação	-0,097*	-0,076	-0,053
	Obter novos financiamentos	0,188*	0,059	0,099*
Total	0,235*	0,189*	0,174*	
Só uso não instrumental		-0,170*	-0,208*	-0,076
Instrumental + Não instrumental		0,260*	0,229*	0,186*
Total USO		0,357*	0,303*	0,215*
Não uso		-0,357*	-0,307*	-0,215*

Fonte: Inquérito por questionário.

Nota: * A correlação é significativa no nível 0,05 (bilateral).

Mas, se considerarmos apenas usos estritamente instrumentais e usos estritamente conceituais e/ou simbólicos, as tendências invertem-se. Com efeito, o

¹²⁶ Para avaliar se a abrangência da participação de *stakeholders* na avaliação (número de grupos de *stakeholders* internos/externos envolvidos) se correlaciona com a de adesão a cada um dos usos/tipos de usos (não/sim) ao recorreu-se ao coeficiente Eta.

¹²⁷ Relembro que a categoria “Total usos instrumentais” abrange utilização só instrumental e utilização instrumental e não instrumental.

crescimento dos níveis de participação das partes interessadas é acompanhando da retração nos usos apenas instrumentais ou apenas não instrumentais, **orientando as organizações para uma maior amplitude e hibridez nos usos dos exercícios avaliativos**. Este efeito é particularmente saliente nas associações entre participação interna e uso não instrumental e entre participação externa e uso instrumental. Ou seja, enquanto a **diversificação do envolvimento de stakeholders externos tende a dissuadir usos estritamente instrumentais**, o **reforço da participação de stakeholders internos tende a desaconselhar usos estritamente não instrumentais**.

4.5.3. Compreendendo os usos da avaliação: entre a institucionalização e a estratégia

No estudo de Eckerd e Moulton (2011) é testada a hipótese segundo a qual enquanto a adoção de determinados métodos de avaliação é largamente definida pelo ambiente das organizações (refletindo a institucionalização), os usos da avaliação são particularmente alinhados com os papéis das organizações (refletindo estratégia). O autor e a autora concluem que ao passo que as pressões institucionais constituem preditores consistentes da adoção de certos tipos de práticas de avaliação, os usos da avaliação variam sobretudo consoante as funções e estratégias específicas das organizações. Vejamos se, neste estudo, as evidências suportam ou não esta conclusão.

As correlações entre usos da avaliação e variáveis que descrevem indutores da avaliação associados às fontes de receitas – “Fazemos avaliação porque os financiadores exigem”; “Fazemos avaliação para nos promovermos perante financiadores e atrair recursos”; “Os nossos financiadores estão muito interessados na avaliação das intervenções” – deixam transparecer a influência das forças externas implicadas nas relações de financiamento nos usos que se fazem da avaliação (vide tabela 10, anexo 5). É visível a associação dos usos exclusivamente não instrumentais (e também do não uso) com a adoção das práticas de avaliação por imposição das entidades financiadoras (“porque os financiadores exigem”), enquanto os usos instrumentais tendem a dissociar-se das pressões do financiamento e, quando apenas instrumentais, até do interesse das entidades financiadoras na avaliação. No caso dos usos (exclusivamente) simbólicos e/ou concetuais, além do padrão mais condicente com os pressupostos da teoria da agência, sobressai

também um arquétipo com maior proximidade aos desígnios da **teoria da dependência**, que associa este tipo de uso ao interesse dos financiadores e à possibilidade de **assegurar recursos e reduzir a incerteza**.

Não obstante, a análise da influência do contexto de financiamento nos usos da avaliação, ponderando os efeitos de variáveis de caracterização da estrutura que descrevem atributos internos das organizações, como seja a dimensão e a antiguidade, vem moderar o peso desta influência que estes resultados parecem indicar.

Procurou-se construir um modelo geral explicativo dos diversos usos e tipos de uso que integrasse não apenas dinâmicas ambientais, mas também fatores associados estrutura e papéis internos e que relativos à estrutura organizativa (tal como se procedeu para o tipo de processos de avaliação adotados)¹²⁸. O quadro que se segue sintetiza os resultados da análise multivariadas, sendo apenas apresentadas as correspondendo às variáveis influentes nos modelos de regressão (*p-value* > 0,005).

Quadro 14. Usos da avaliação: Perfis de OTS

		DIMENSÃO	TIPO ENTIDADE	FONTE DE RECEITAS	FUNÇÕES	IDADE	TOTAL*
Usos instrumentais	Alocação recursos	Grandes	IPSS/Outras /Coletiv.		Serviços/ Mistas		3
	Decisões sobre RH	Grandes/ Médias			Serviços/ Mistas		2
	Rever estratégias	Grandes/ Médias	(- ADL)			Recente/ Antiga	3
	Rever projetos		ADL/ONG		Expressão/ Mistas	Recente/ Intermédia	3
	Total	Grandes			Mistas		2
	Uso só instrumental						0

¹²⁸ A análise realizada mediante o recurso à regressão binominal materializou-se na criação de **um modelo para cada um dos usos e tipos de usos (15 modelos)**. Como já foi dito, este tipo de análise permite-nos considerar todos os fatores em simultâneo, dando-nos uma perspetiva dinâmica de interinfluência.

Usos não instrumentais	Reportar financiadores		ONG	Públicas		Intermédia	3
	Reportar órgãos sociais						0
	Obter certificação		ONG/ADL/ IPSS	Públicas (projetos)			2
	Rel. públicas e divulg.	Médias			Expressão	Recente	3
	Obter novos financiamentos	Pequenas/ Médias		Subsídios		Recente/ Intermédia	3
	Total					(- Antigas)	1
Uso só não instrumental	Pequenas				(- Mistras)	2	
Uso misto					Mistas	Recentes	2
Não uso							0
TOTAL**	7	5	3	7	7		

Fonte: Inquérito por questionário.

Notas: * N.º total de variáveis independentes (dimensão, tipo de entidades, ...) que revelam poder preditivo na análise multivariada.

** N.º total de casos em que a variável independente revela poder preditivo do comportamento da variável dependente na análise multivariada.

Uma primeira leitura e global do quadro é reveladora da relativa **menor influência das pressões institucionais associadas às fontes de financiamento** quando comparada com a exercida por fatores de contingência internos como sejam a antiguidade, a dimensão e as funções/papéis organizacionais. É também mais modesta quando comparada com o peso que tem no tipo de práticas de avaliação adotadas, em que o modelo de financiamento está no topo das variáveis influentes da maioria dos processos de gestão e avaliação elencados, como vimos anteriormente (vide quadro 9 do ponto 4.3.4)¹²⁹. Vemos que, em termos globais, as **variáveis ambientais**, particularmente as associadas ao financiamento, são menos frequentemente correlacionadas com os usos da avaliação quando comparados com as práticas adotadas. Isto significa que, mesmo quando a adoção de práticas de avaliação é primeiramente influenciada por fatores externos, as OTS procuram beneficiar do uso da avaliação no sentido de promover a eficácia, sugerindo que **os usos da avaliação são menos institucionalizados do que as práticas adotadas**, na linha do observado por Eckerd e Moulton (2011).

¹²⁹ É de notar, em todo o caso, que o poder explicativo das regressões realizadas para os usos é genericamente mais modesto que o verificado para os tipos de práticas implementadas, o que significa que magnitude da influência das características organizacionais é globalmente menor nos usos que se fazem da avaliação do que no tipo de práticas adotadas.

De resto, alguns atributos internos que tinham na análise do tipo de processos avaliativos uma relevância marginal no conjunto dos fatores considerados, assumem aqui uma influência mais marcada. É o caso da antiguidade, que assume um papel particularmente relevante na compreensão da adesão aos usos não instrumentais da avaliação, a que tendem a dissociar-se as organizações de constituição mais antiga, eventualmente menos preocupadas em demonstrar legitimidade. Mas é também o caso e do tipo de funções que, por oposição, assume uma posição proeminente na explicação da adesão a usos instrumentais (e mistos). A forte influência das principais funções organizacionais nos usos da avaliação sugere que os usos se encontram mais alinhados estrategicamente com os papéis organizacionais dos que os métodos de avaliação adotados, conforme defendido no trabalho de Eckerd e Moulton (2011), e está em linha com os pressupostos da **teoria da contingência** em cujo enquadramento a avaliação é usada como mecanismo facilitador do alinhamento entre diversos papéis e expectativas externas. A heterogeneidade de papéis do terceiro setor, incluindo a prestação de serviços, a criação de capital social, a cidadania, a inovação e a expressão de valores e advocacia política, compõe a vantagem destas organizações e a mais-valia do seu contributo para a sociedade.

Vemos que, em termos gerais, a prevalência de **funções de expressão** de valores, interesses e crenças culturais, espirituais, profissionais ou políticas (e com características semelhantes em termos de dimensão, antiguidade e perfil de receitas) **é decisiva para a compreensão da dissociação a usos com caráter material, e da adesão a usos de caráter imaterial enquadreadores de estratégias de aproximação aos públicos**. Já a acumulação de funções de expressão com funções de serviço é particularmente relevante não só para fundamentar o uso da avaliação para auxiliar a tomada de decisão quotidiana e estratégica, mas também a acumulação destes propósitos materiais a outros de caráter imaterial (usos mistos).

Análises anteriores dão suporte à ideia da maior associação das funções de **expressão a usos não instrumentais** da avaliação (Eckerd e Moulton, 2011; Højlund, 2014). No seu estudo empírico, Eckerd e Moulton (2011:114) observam que, enquanto o aumento da ênfase na prestação de serviços e na inovação está associada ao uso instrumental da

avaliação, o aumento da ênfase na advocacia política ou na criação de capital social se associa ao aumento nos usos simbólicos. Højlund (2014) outorga esta tendência às possibilidades de medição oferecidas pelas funções desempenhadas pelas OTS, asseverando que quando os papéis das organizações são difíceis de medir – como é o caso da defesa de direitos e interesses profissionais, culturais, políticos ou sociais – as organizações tendem a usar a informação simbolicamente. Sendo válida a ideia do desafio acrescido que se coloca à medição nas entidades com funções expressivas, também refletida nas análises de Ebrahim e Rangan (2010, 2014), a aceção de Højlund pressupõe uma visão estrita da eficácia organizacional, que não acolhe a possibilidade de, para algumas OTS, a eficácia estar diretamente associada aos usos simbólicos da informação sobre o desempenho.

Como notam Eckerd e Moulton (2011), embora os usos instrumentais sejam frequentemente promovidos, os outros usos da avaliação são igualmente – se não mais – importantes para muitas organizações do terceiro setor, como, aliás, tem sido demonstrado em diversos estudos (Carman, 2008, Ellis e Gregory, 2008; Ógáin, Lumley e Pritchard, 2012). **Para algumas OTS, como grupos de advocacia política, a eficácia pode estar diretamente relacionada com os usos simbólicos** (DiMaggio, 2001)¹³⁰. Conforme asseveram Eckerd e Moulton (2011), os propósitos de demonstração podem, sem dúvida, reforçar a eficácia da organização, quando a eficácia é considerada não apenas como a capacidade de atingir objetivos, mas também como a capacidade de atrair recursos, de sobrevivência organizacional e de adquirir legitimidade a longo prazo perante as partes interessadas relevantes. **Donde, as razões para os usos não instrumentais não se confinam, necessariamente, à intangibilidade dos papéis das organizações**, mas poderão enquadrar **estratégias de reforço da eficácia** mediante a aquisição de **legitimidade** perante as partes interessadas.

Apesar da relevância da função concetual ou mesmo simbólica da avaliação para a eficácia das organizações, a função instrumental da avaliação é crítica para a

¹³⁰ Por exemplo, a legitimidade sinalizada pelo cumprimento de normas externas pode resultar num maior apoio político para as causas da organização e, posteriormente, para o seu funcionamento, por meio de subsídios ou contratos que possam resultar do processo político (Eckerd e Moulton, 2011: 101). A adoção de técnicas de avaliação que não permitem que tais organizações demonstrem legitimidade pode ser inútil (DiMaggio, 2001).

aprendizagem nas organizações. Com efeito, o **estudo demonstra que os propósitos de legitimidade apenas se revelam vantajosos quando, além do uso não instrumental, os resultados da avaliação auxiliam a ação direta.** Conforme vimos anteriormente, os usos exclusivamente não instrumentais associam-se às representações mais negativas sobre avaliação – à percepção da avaliação como desvio do foco da intervenção, como desperdício/inutilidade e como fardo imposto pelo financiamento – ao passo que a combinação de usos instrumentais e não instrumentais se revela mais profícua em termos de internalização das práticas de avaliação.

4.6. Síntese: a prevalência das lógicas managerialistas e tecnocráticas de medição do desempenho

Um resultado que compõe o pano de fundo de toda a análise respeita ao conceito de avaliação. Existem no seio das organizações perspectivas muito diversas sobre o conceito, sendo amplo o leque de práticas identificadas como sendo práticas de avaliação: desde práticas informais e reativas até exercícios mais sistemáticos e pró-ativos, sejam especificamente avaliativos, sejam de gestão, mais genericamente. A amplitude do conceito encontra-se refletiva na diversidade de enquadramentos das atividades de avaliação nas organizações, sendo particularmente frequente a condução de exercícios de avaliação ou medição do desempenho no âmbito do planeamento operacional (planos anuais), estratégico, da implementação de projetos e da gestão da qualidade. De resto, a generalidade das entidades auscultadas no estudo declara fazer pelo menos alguma monitorização e avaliação das suas atividades. Esta tendência alinha-se com a subscrição efusiva da retórica da avaliação, mas que não é totalmente condicente com as práticas descritas e com o destaque que lhe é atribuído. Com efeito, apesar da desejabilidade e relevância da avaliação nos discursos, a reflexão sobre as intervenções, sobre os respetivos processos e resultados, é claramente um domínio não prioritário de ação das organizações.

A profusão concetual sobre avaliação que é identificada no terreno é coerente com a encontrada na revisão da literatura, onde encontramos entendimentos sobre avaliação muito diversos consoante a área do conhecimento em que se inscrevem as discussões (teóricas ou empíricas). As OTS oferecem um enquadramento excecional para o

cruzamento de diferentes lógicas e linguagens sobre avaliação produzidas nos diversos campos de estudo, já que o terceiro setor tem sido descrito como híbrido, em virtude de ocupar uma posição intermédia entre o mercado, o Estado e a comunidade. Ora, a discussão sobre avaliação que se faz no terceiro setor é permeável às diversas racionalidades geradas nestes três polos, pelo que as lógicas de avaliação das organizações do terceiro setor se situarão num campo híbrido entre diversos tipos de discurso associados a cada um deles, nomeadamente dos discursos científico, da gestão e da sociedade civil. A análise revelou a secundarização da linguagem que reflete valores associativos face à linguagem alicerçada nos outros domínios discursivos, sobretudo a originada na cultura gestonária. Embora se tenham observado sinais de combinação (e até de reforço mútuo) dos diversos tipos de linguagem, expressos pela “interlinguagem”, a prevalência da orientação discursiva com raízes na gestão evidencia a colonização da linguagem de génese associativa pela linguagem gestonária, parecendo subordinar-se o carácter “expressivo” das organizações (isto é, de expressão de valores e ideais de serviço), ao seu carácter “instrumental” (orientação para o desempenho e para os resultados). A preocupação que daqui decorre é que a predominância do discurso gestonário (e científico), ao promover sua orientação instrumental, possa estar a excluir as vozes do terceiro setor, perdendo de vista os seus valores constitutivos.

A análise dos tipos de práticas e dos modelos de avaliação prevalecentes nas organizações, em termos do seu conteúdo e desenho metodológico, vem dar algum fundamento a esta preocupação. Vemos que, tal como propus na minha primeira hipótese, as lógicas managerialistas e tecnocráticas de medição do desempenho são prevalecentes nas organizações. A avaliação, sobretudo a conduzida ao nível da organização como um todo, tende a assumir no estudo feições próximas das práticas de controlo de gestão do setor empresarial, em que as análises mais compreensivas e contextuais tendem a ceder lugar a uma lógica mais alinhada com os pressupostos da medição do desempenho, em articulação com sistemas de gestão mais amplos, como a gestão estratégica e da qualidade. Esta lógica tem permeado também a avaliação conduzida no âmbito dos programas comunitários em que informação requerida se destina essencialmente a alimentar

indicadores quantificados de desempenho (de curto prazo) dos programas (é o caso do POPH).

A caracterização das práticas de avaliação conduzidas nas organizações no que se refere ao seu conteúdo e desenho metodológico apontam, em termos genéricos, para a prevalência de perspectivas de avaliação que são enquadráveis na primeira e segunda gerações da avaliação, que Guba e Lincoln (1989) nomeiam de fases da “medição” e da “descrição”. Os procedimentos de avaliação dominantes, seja a nível dos projetos seja a nível organizacional, envolvem a recolha de dados quantitativos e mensuráveis, em que a medição dos fenómenos é frequentemente complementada com a sua descrição, mas sem apresentar apreciações críticas e sustentadas sobre o valor e mérito das intervenções. As apreciações qualitativas e críticas, quando são incluídas, tendem a alicerçar-se em apreciações rápidas e intuitivas sobre a forma como as atividades decorreram e os resultados que foram alcançados, sem recurso a métodos sistemáticos de produção de informação.

Estas perspectivas de medição e descrição articulam-se com o centramento dos exercícios avaliativos na resposta a objetivos. A abordagem por objetivos, radicada no paradigma positivista, constitui, pois, a perspectiva dominante nas organizações estudadas, nos dois níveis de análise – de projetos e organizacional. A nível dos projetos, a adoção do modelo por objetivos por parte das organizações reflete, pelo menos em parte, a lógica da generalidade dos programas de financiamento nacionais, comunitários e internacionais. A nível organizacional, este modelo é predominante não apenas nos exercícios avaliativos conduzidos no âmbito do planeamento estratégico e operacional, mas também no âmbito da certificação da qualidade, processos fortemente arraigados na filosofia de gestão por objetivos. A simplicidade e produtividade da avaliação por objetivos, assim como a facilidade em legitimar os seus resultados (Stockmann, 2011), torna-a atrativa quer para entidades financiadoras quer para OTS. Apesar das suas vantagens, nomeadamente por induzir a reflexão sobre as finalidades e metas das intervenções ou das organizações como um todo, a abordagem por objetivos (de tradição positivista, racional) tende a ser inflexível e linear, induzindo o fenómeno de visão de túnel (*ibidem*). Ao omitir informação de contexto e sobre os efeitos não previstos ou fora do âmbito dos objetivos fixados pelas

organizações, o enfoque nos objetivos negligencia perspectivas particularmente relevantes em entidades com uma missão social e que atuam em ambientes complexos e incertos, como são as OTS.

A conceção de eficácia centrada no alcance de objetivos, prevalente no estudo, articula-se com a centralidade da abordagem de processos, que faz incidir as análises nos recursos, atividades e realizações/*outputs*, conforme previsto na alínea a) da minha primeira hipótese. Os esforços mais significativos das organizações são dedicados à produção de informação sobre níveis de execução financeira e física e sobre os resultados de tipo mais imediato que a atividade das organizações produziu (e.g, número de pessoas formadas, grau de satisfação dos/as utilizadores/as). O foco tende a ser colocado nos processos internos das organizações, incluindo a forma como os serviços são prestados e as práticas de gestão, i.e., naquilo a que Osborne (1992) se refere como “qualidade do serviço”. A ausência da perspectiva ancorada nos benefícios sociais espelha, em certa medida, a própria orientação dos requisitos associados ao financiamento e aos sistemas de certificação/acreditação de produtos/serviços das OTS, ainda limitados no propósito de direcionar os processos de monitorização e avaliação das intervenções para as mudanças. Uma vez que fornece a quem avalia uma estrutura para analisar os principais processos que influenciam o desempenho, a abordagem dos processos tem a vantagem de potenciar a compreensão das alavancas do processo de criação de valor para os diferentes *stakeholders* – essencial para compreender o para explicar “como” e o “porquê” a mudança acontece. Todavia, é incipiente no propósito de informar sobre as mudanças em si, i.e., sobre os efeitos da intervenção na qualidade de vida das pessoas – dimensão que estabelece ligação com a missão das organizações e com o valor social criado. Ou seja, além de ser limitado no propósito de avaliar o cumprimento da missão (quando ela remete para efeitos ou impactos) (i.e., ao nível organizacional), é insuficiente para aferir sobre o valor social criado para a sociedade (i.e., ao nível do sistema social). Depois, ao secundarizarem os efeitos/mudanças, nas suas análises avaliativas, as organizações tenderão a concentrar-se menos nas pessoas beneficiárias e outros *stakeholders*, já que o foco nas realizações/*outputs* remete para as atividades e não para quem delas beneficia.

Encontramos aqui, pois, elementos para a compreensão das barreiras à afirmação de modelos de avaliação participativos nas OTS estudadas.

As abordagens construtivistas ou transformativas à eficácia (reputacionais/"emergentes"/de múltiplos constituintes), nomeadamente assentes em metodologias de avaliação participativas/colaborativas, encontram-se, efetivamente, pouco disseminadas nas organizações. O estudo denuncia, grosso modo, a dificuldade das organizações em promover uma avaliação participativa – que ultrapassasse o nível superficial da consulta ao público-alvo. São raros os casos em que as iniciativas de avaliação se abrem à negociação entre as diferentes partes interessadas sobre o que vai ser monitorizado e avaliado, sobre as condições em que os dados são recolhidos e analisados, o realmente significam, ou sobre a forma como os resultados serão partilhados e usados. Embora seja generalizado o reconhecimento da relevância do envolvimento das diversas partes interessadas na avaliação, ele aparece situar-se mais no plano discursivo dos desejos e aspirações do que em práticas abrangentes, estruturadas e aprofundadas de participação. O facto de entre as práticas mais mobilizadoras da participação nas OTS se encontrarem as dinamizadas por iniciativa de atores externos às OTS (e.g., auditorias) é, de resto, reveladora da dificuldade das organizações em promover a participação alargada nos processos de gestão e avaliação. Por outro lado, é destacado o contributo da condução de exercícios de avaliação estruturados, sobretudo à escala dos projetos, para a abertura da avaliação à participação de atores externos às organizações. As experiências de avaliação participativas mais inovadoras e aprofundadas relatadas no estudo são conduzidas no âmbito de projetos (nomeadamente do PIC EQUAL) ou passaram a ser integradas na rotina das organizações depois de serem experimentadas em projetos, ainda que as práticas mais regulares e institucionalizadas (mas menos aprofundadas e mais simbólicas) de envolvimento de *stakeholders* decorram dos processos de certificação/acreditação para o domínio da formação ou na área da qualidade. Perante a fragilidade dos mecanismos de participação nas organizações estudadas não é de estranhar a exiguidade das mudanças cognitivas ou comportamentais decorrentes da "exposição" à atividade de avaliação, para além do âmbito restrito dos profissionais e chefias diretamente implicados na dinamização das atividades de monitorização e avaliação. Não obstante o aparente parco contributo dos

processos avaliativos conduzidos pelas entidades, que Patton (1997) cunhou de uso de processo, o estudo mostra que, em termos gerais, o envolvimento das partes interessadas aumenta a probabilidade de os seus resultados serem utilizados, reiterando um argumento fortemente invocado para a adesão a abordagens participativas (Fine, Colette e Coghlan, 2000; Cousins, Goh e Elliott, 2008; Fleischer e Christie, 2009; Johnson *et al.*, 2009).

A relevância deste e de outros fundamentos que têm motivado a adoção de abordagens de avaliação pluralistas (nomeadamente as preocupações com a justiça social, inclusão democrática, empowerment e emancipação, e a centralidade do contexto e da inclusão de múltiplos participantes na produção do conhecimento), e a sua particular afinidade com os valores e funções das OTS, justificariam um maior investimento neste tipo de abordagem.

Capítulo 5. Capacidade *para fazer avaliação*: contextos e competências

A capacidade de avaliação refere-se à aptidão das entidades para conduzir, usar e aprender continuamente a partir dos processos de avaliação (Preskill e Torres, 2000; Preskill e Boyle, 2008). Algumas análises têm observado que a capacidade da gestão das OTS para responder aos desígnios da avaliação, designadamente os estabelecidos pelas entidades financiadoras, não tem acompanhado o ritmo da crescente procura deste tipo de informação (Newcomer, 2004). Esta perceção tem resultado num crescente interesse entre investigadores/as no sentido de encontrar formas de melhorar a capacidade de avaliação destas entidades, mediante atenção dedicada à implementação de iniciativas de capacitação em avaliação (*evaluation capacity building/evaluation capacity development*) ou de institucionalização da avaliação. Muito deste interesse tem-se materializado no desenvolvimento de modelos explicativos para descrever e ilustrar as condições necessárias para alicerçar a avaliação nas organizações (Stufflebeam, 2002; Valkov e King, 2005; Bourgeois e Cousins, 2008; Powell e Boyd, 2008; Preskill e Boyle, 2008). Recentemente, Bourgeois e Cousins (2013) propuseram um quadro que ilustra a complexidade e amplitude das atividades associadas ao desenvolvimento da capacidade avaliativa (vide quadro 15). As autoras concebem a capacitação em avaliação enquanto processo pelo qual as organizações desenvolvem competências para conduzir e utilizar os resultados da avaliação num conjunto de seis dimensões. A capacidade para implementar avaliações envolve pessoal e recursos organizacionais mais amplos, bem como a mobilização de apoios internos e externos para planear e implementar avaliações; enquanto a capacidade para usar a avaliação requer uma cultura organizacional que promova a participação nas atividades de avaliação, a integração da avaliação na tomada de decisão organizacional, assim como a aplicação das conclusões da avaliação (Bourgeois e Cousins, 2008).

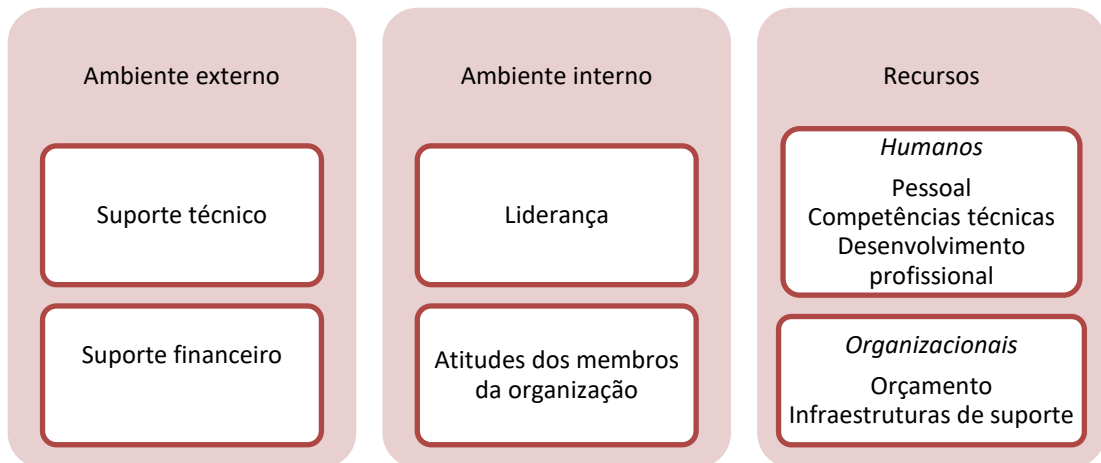
Quadro 15. Dimensões da capacidade de avaliação das organizações

Capacidade para fazer		
Recursos humanos: <ul style="list-style-type: none"> • Pessoal • Lógica e competências de avaliação • Competências de comunicação e relacionamento interpessoal • Desenvolvimento profissional • Liderança 	Recursos organizacionais: <ul style="list-style-type: none"> • Orçamento • Recolha de dados contínua • Infraestrutura organizacional 	Planeamento e atividades de avaliação: <ul style="list-style-type: none"> • Plano de avaliação • Recurso a consultoria • Partilha de informação • Apoios externos • Ligações organizacionais
Capacidade para usar		
Literacia em avaliação: <ul style="list-style-type: none"> • Envolvimento da avaliação • Gestão orientada para resultados 	Processo de decisão organizacional: <ul style="list-style-type: none"> • Processos de decisão • Apoio à decisão 	Benefícios de aprendizagem: <ul style="list-style-type: none"> • Uso instrumental ou concetual • Uso de processo

Fonte: *Adaptado* de Bourgeois, Isabelle; Cousins, J. Bradley (2013) "Understanding dimensions of organizational evaluation capacity" *American Journal of Evaluation*. 34(3), 299–319.

Neste capítulo analisam-se algumas das **dimensões associadas à capacidade para fazer avaliação**, associada ao conhecimento e às competências necessários para conduzir ou administrar atividades de avaliação (vide figura 6). Além de questões associadas aos **recursos** disponíveis para avaliação, quer **humanos** (pessoal, competências técnicas, desenvolvimento profissional), quer **organizacionais** mais amplos (orçamento, fontes de financiamento, infraestruturas de suporte), serão abordadas algumas dimensões associadas ao **contexto organizacional interno** (liderança e atitudes dos membros da organização) e **externo** (suporte técnico e financeiro de entidades financiadoras, estruturas representativas e outros atores relevantes), dimensões que têm sido particularmente associadas à capacidade para fazer avaliação (Preskill e Torres, 1999; Botcheva, White e Huffman, 2002; Volkov e King, 2007; Alaimo, 2008; Bourgeois e Cousins, 2008; Hole e Patterson, 2008; Preskill e Boyd, 2008).

Figura 6. Dimensões analíticas da capacidade *para fazer avaliação*



Fonte: Elaboração da autora.

5.1. Estruturas organizacionais de suporte à avaliação

5.1.1. Recursos humanos

Apesar de a generalidade das organizações afirmar fazer algum tipo de monitorização e avaliação das suas atividades e de estes processos serem assegurados, na maioria dos casos, por **pessoal interno**¹³¹ (com ou sem envolvimento nas atividades a serem avaliadas), ela é **pouco profissionalizada**, já que na maioria dos casos os/as responsáveis pela avaliação não são avaliadores/as profissionais nem têm competências específicas em avaliação. Além disso, tal como observou Mesquita (2015) no seu estudo sobre cultura de avaliação no setor da cooperação para o desenvolvimento, são **raras as situações em que existe alguém da equipa de trabalho dedicada à avaliação em tempo integral**. No entender de responsáveis de OTS entrevistados/as, ter alguém a trabalhar em avaliação a tempo inteiro, além de não ser viável, dada a fraca capacidade financeira deste tipo de entidades, não parece apresentar vantagens em termos de retorno do investimento, como se pode ver nos seguintes excertos de entrevistas:

Tem a ver com as limitações ao nível da estrutura de RH, o tempo de trabalho que podem dedicar. Porque muitas vezes não há uma pessoa a trabalhar só na

¹³¹ Em termos de intervenientes no processo de avaliação, a modalidade mais frequente é a autoavaliação: em perto de 90% das organizações inquiridas, a avaliação é dinamizada pelas próprias pessoas implicadas na atividade em análise. A avaliação interna independente (que é conduzida por pessoas da organização sem envolvimento na atividade a ser avaliada) é bem menos frequente, abrangendo menos de um quarto das entidades.

parte da avaliação, não pode haver. Não há capacidade financeira para ter uma pessoa só a trabalhar nisso. É parte do trabalho normal que os RH fazem de toda a gestão dos projetos e organizações. (...) E como não é uma área em que as organizações vejam muito retorno, não dedicam tempo suficiente. (Entrevista com representante de organização de cúpula – 43)

Em termos de avaliação, se os programas não são muito exigentes, se não nos financiam, nós também não temos capacidade para fazer um investimento tão grande a esse nível porque acabamos por não ter retorno. (...) (Entrevista com dirigente de ADL – 20)

Apesar de os agentes da avaliação serem sobretudo internos, a maioria das organizações (cerca de 58% das entidades inquiridas) **não conta, internamente, com pessoas com competências específicas em avaliação**. A análise das entrevistas confirma e reforça esta observação. Na grande maioria das organizações entrevistadas existem pessoas com competências técnicas mais genéricas que lhes permitem responder minimamente aos requisitos de monitorização e avaliação das atividades, mas que raramente lhes conferem competências para conduzir exercícios de avaliação estruturados ou em profundidade.

Não temos aqui a possibilidade de criar áreas de especialização. Dizer que temos uma pessoa aqui só a fazer avaliação, nós não temos. Temos pessoas que vão frequentando algumas ações de formação no sentido de ficar minimamente capacitadas, ou com competências para dar resposta àquilo que são as exigências dos financiadores. (Entrevista com coordenação técnica de ADL - 19)

Nesta linha, acompanhando os diversos enquadramentos (projetos, gestão da qualidade, planeamento estratégico, planeamento operacional) e o enorme leque de práticas que as organizações identificam como sendo práticas de avaliação (avaliação de programas/projetos, análises de satisfação, diagnóstico organizacional, medição dos desempenhos, avaliação das aprendizagens etc.), encontramos uma grande **diversidade**, quer de **perfis funcionais de agentes** responsáveis pela condução destes exercícios das organizações, quer de **competências** que as organizações identificam como **avaliativas**. Assim, quando a avaliação é enquadrada em processos de planeamento anual, os principais responsáveis pela avaliação são os responsáveis setoriais/de cada área ou serviço; quando se enquadra em processo e planeamento a médio ou longo prazo, a responsabilidade é partilhada pela coordenação/direção executiva e responsáveis setoriais; quando se enquadra nos processos de gestão da qualidade é liderada pelos responsáveis dos departamento da qualidade e, finalmente, quando a avaliação se enquadra em projetos, é dinamizada pela coordenação/equipa técnica a eles afeta. Ao dizerem que existe na

organização pessoal com *expertise* em avaliação (o que nas entrevistas foi pouco frequente), os/as representantes das OTS tendem a falar de **aptidões e atribuições diversas**, dependendo do enquadramento dos processos de avaliação e dos seus agentes.

O recurso à **avaliação externa**, realizada por especialistas sem vínculo à organização, é prática comum em apenas **um quarto das entidades** inquiridas. A leitura deste valor de adesão que, ainda que bem menos enfático que o relativo à avaliação interna, abrange uma proporção significativa de entidades, deve ser feita à luz de uma conceção abrangente sobre o tipo de serviço prestado no âmbito de uma avaliação externa. As entrevistas evidenciam que, tal como acontece relativamente à avaliação interna, o tipo de serviço prestado no âmbito das consultorias que os atores das organizações identificam como avaliativas é muito diverso, podendo compreender desde a supervisão da **implementação dos sistemas de qualidade** no âmbito da certificação da qualidade (auditorias internas feitas por consultores externos e auditorias externas feitas por entidades certificadoras), até à **assessoria/supervisão científica** na condução das atividades (candidaturas, conceção, implementação e avaliação), à realização de **estudos de avaliação** de iniciativas/projetos específicos por parte de profissionais externos, ou às ações de **monitoria** da iniciativa de entidades financiadoras. Eis algumas situações relatadas em entidades auscultadas em entrevista:

Temos uma equipa externa que faz a supervisão disto do ponto de vista dos mecanismos da qualidade, que obviamente incluem a monitorização e avaliação. (...) Fizemos um concurso público e é uma empresa cujo responsável é especialista nisto e que é responsável pela avaliação da qualidade. (Entrevista com dirigente de ONG – 16)

Inicialmente nós recorriamos muito ao apoio de consultores que sabiam trabalhar as questões da avaliação, porque o EQUAL previa. (...) Entretanto nunca mais trabalhámos... Isso é um pouco aquilo que foi a nova configuração do QREN, que deixou de privilegiar o financiamento desse tipo de recursos e nós deixamos de recorrer. (Entrevista com dirigente de ADL – 20)

A opinião sobre a avaliação externa é genericamente **positiva**: para mais de 70% das organizações inquiridas – em especial para as IPSS –, a experiência de colaboração com avaliadores/as externos/as é boa ou excelente. Apenas 1% a classifica como má ou péssima (e menos de um terço classifica-a como razoável). Nas entrevistas, as apreciações sobre a consultoria externa também são genericamente positivas, ainda que tenham lugar algumas visões críticas – não só de representantes de OTS mas também de profissionais de avaliação entrevistados/as - que incidem sobre a **falta de qualidade** de alguns trabalhos e da pouca

atenção prestada aos **impactos**, por um lado, e às **especificidades** das intervenções e das organizações visadas, por outro:

Há coisas muito díspares do ponto de vista da qualidade técnica e da confiança que transmitem. Quando eu leio certos relatórios de avaliação (...) tenho a consciência que por vezes tanto faz fazer bem, mal ou mais ou menos, porque vai ser tudo aceite da mesma forma, e como uma coisa fantástica. (...) E acabam por não retirar a mais-valia porque não foram exigentes no caderno de encargos (Entrevista com consultor/profissional de avaliação – 37)

Em termos muito gerais, para a área internacional, o que acontece muitas vezes é quem faz avaliação de projetos internacionais está muito focado nas questões de como é que a organização que foi financiada consegue provar que as atividades foram feitas. E, portanto, uma muito menor atenção a questões de impacto. (Entrevista com dirigente de ONG – 6)

De acordo com as pessoas entrevistadas, o pessoal interno e o pessoal externo trazem **diferentes “valores”** para a avaliação: enquanto o pessoal interno carrega o conhecimento aprofundado sobre as intervenções e sobre as próprias organizações (e suas especificidades), o pessoal externo garante *expertise*, objetividade, rigor, “distanciamento” e independência na apreciação do trabalho desenvolvido.

Acho que é positivo e traz aqui mais-valia, porque é um olhar de fora. (...) Porque são pessoas que não estão diretamente envolvidas no trabalho e conseguem ter uma perceção diferente daquilo que está a ser feito e que nos podem dar achegas importantes. (Entrevista com coordenação técnica de IPSS – 11)

Um dos benefícios dos aportes distintos da consultoria externa mais valorizados pelas organizações é a **maior credibilidade** dos exercícios de avaliação, que acabam por ser mais potenciadores e influentes em termos de certificação e legitimação da ação das organizações: “Esta componente externa é importante porque nos dá um maior rigor de ter uma assessoria externa que nos diz e que vai certificar que o que nós estamos a fazer é bem feito” (Entrevista com direção técnica de ADL – 7).

Algumas pessoas responsáveis de OTS entrevistadas sugerem que a utilidade das avaliações resulta da **combinação de visões** internas e externas sobre as intervenções, a qual permite agregar as vantagens e mitigar as limitações de cada uma delas:

Quando alguém vem aqui fazer uma avaliação tem que conjugar a avaliação externa com a autoavaliação, no sentido de ela própria se integrar e ser sentida como uma pessoa que também faz parte. Porque as pessoas retraem-se. A maioria das pessoas não vê a avaliação como um processo de melhoria do projeto mas sim como uma arguição do próprio. (Entrevista com técnico de ONG – 14)

Por fim, encontramos uma visão que pudemos classificar de **emancipatória** relativamente ao papel da consultoria externa em avaliação. Nestes casos, o papel da

consultoria é formativo, no sentido de capacitar os atores internos para condução dos exercícios de avaliação. A estratégia consiste em aceder a serviços especializados com o objetivo de dotar a equipa técnica de ferramentas que permita à organização conduzir internamente a avaliação. O testemunho que se segue é exemplar:

Nós temos trabalhado e temos trazido avaliadores, mas não para a avaliação das intervenções, mas para melhorar as nossas perspetivas de avaliação, no sentido de reforçar competências, de mostrar como é que a organização pode fazer a avaliação. Mas nunca fazemos avaliação externa. Quer dizer, somos monitorizados. Como trabalhamos na área da cooperação há sempre avaliadores que avaliam as nossas intervenções e dão uma perspetiva bastante interessante. Mas internamente para as intervenções que desenvolvemos na nossa região temos utilizado a avaliação feita pela equipa. (Entrevista com coordenação técnica de ADL – 26)

Em estudos conduzidos em contexto britânico e norte-americano (Fine, Thayler, Coghlan, 2000; Carman, 2007; Ellis e Gregory, 2008; Innovation Network, 2012, 2016; Chung e Tran, 2015), a modalidade de avaliação interna é também a mais frequente nas organizações do terceiro setor¹³², ainda que a avaliação apareça no contexto das organizações destes países como uma **atividade mais especializada**. Por exemplo, o inquérito a OTS norte-americanas da Innovation Network (2012) revela que perto de 20% das entidades conta com uma pessoa interna dedicada à avaliação a tempo inteiro, proporção que sobe para mais de metade em organizações de grande dimensão financeira. Segundo este estudo, em cerca de um quarto das entidades a avaliação é assegurada por avaliadores/as profissionais (internos ou externos). Um outro estudo empírico, conduzido no Reino Unido (Ógáin, Lumley e Pritchard, 2012), observa que mais de um terço das entidades do terceiro setor conta com uma pessoa ou equipa de trabalho cuja principal responsabilidade consiste em avaliar o impacto da sua intervenção.

Ora, no estudo português, **a existência de tal estrutura de avaliação é muito rara**. Além de a maioria das organizações não ter capacidade financeira para integrarem recursos humanos a trabalhar exclusivamente nesta área, **também não parece existir a convicção**

¹³² Na maioria dos estudos referenciados o peso da avaliação externa no total das avaliações conduzidas nas organizações não ultrapassa os 10%.

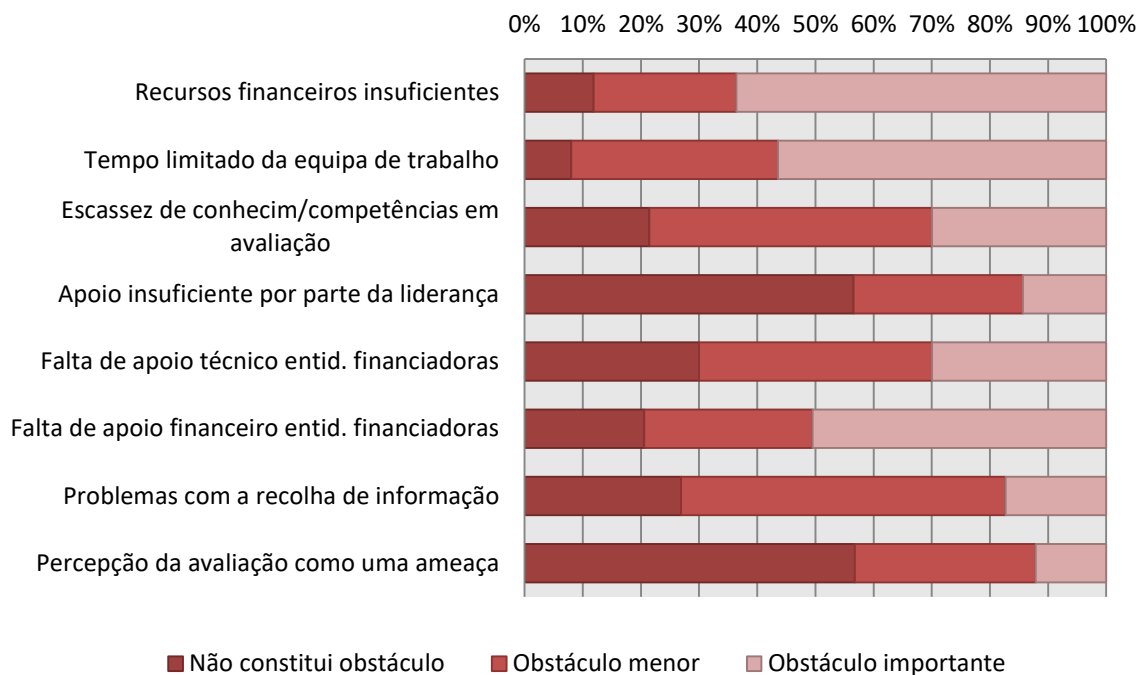
da necessidade e/ou vantagem de ter pessoas a trabalhar exclusivamente nestas funções.

Nas entrevistas encontramos apenas um caso em que na organização (uma IPSS) existe uma técnica cuja principal responsabilidade consiste em “gerir os processos de *accountability* e impacto social” da entidade. Tratava-se de uma técnica ao abrigo de um estágio profissional, com trabalho académico desenvolvido neste domínio, e que propôs à entidade (com a qual já tinha estabelecido ligação no âmbito académico e da participação em redes) desenvolver trabalho nesta área. Ou seja, o impulso para a estruturação do trabalho no domínio da avaliação veio de uma pessoa exterior à organização, que procurou demonstrar a relevância do trabalho neste âmbito. Mas havia também da parte da direção da entidade uma cultura de abertura à condução deste tipo de processos, propiciada pela participação em plataformas de intervenção e em redes de profissionais:

Juntando o trabalho que estava a ser desenvolvido por mim através da [rede de profissionais] ao trabalho que a plataforma X estava a fazer, surgiu a proposta de fazer um estágio que iria dar resposta às necessidades identificadas quer pela [X] quer pela plataforma que já trabalhava com eles. E foi com relativa facilidade que acolheram esta ideia (...), porque era uma coisa que se começava a falar cada vez embora ainda não houvesse muita experiências, principalmente me abordagens mais colaborativas que não o SROI. (Entrevista com técnica de IPSS – 4)

Na generalidade dos casos, o que acontece é que a recolha de dados para produção de relatórios e avaliações é assegurada por pessoal executivo ou técnico não especializado em avaliação, que acumulam com uma série de outras atividades e responsabilidades.

Gráfico 8. Constrangimentos à avaliação



Fonte: Inquérito por questionário.

A **escassez de conhecimentos e competências em avaliação** é considerada um obstáculo à condução de exercícios de avaliação pela maioria das organizações inquiridas, embora menos de um terço considere ser um obstáculo importante, sendo **secundarizado** em relação à escassez de recursos financeiros (vide gráfico 5). Mas se juntarmos à escassez de *expertise* em avaliação os problemas com a **alocação de pessoal**, as limitações associadas aos recursos humanos aparecem no topo das preocupações das entidades. Com efeito, o tempo limitado da equipa de trabalho é um constrangimento apontado por mais de 90% das entidades e considerando um obstáculo importante à monitorização e avaliação das intervenções por mais de metade: “A questão é que a pessoa nem sempre tem condições práticas para fazer este trabalho. São pessoas que têm que dar resposta a muitas solicitações, e não dá para tudo...” (Entrevista com dirigente de OTS – 20)

Além de contarem com colaboradores/as em número insuficiente (Franco, 2015), a gestão de recursos humanos em muitas das OTS é pautada por grandes dificuldades que decorrem da debilidade dos suportes da sustentabilidade do emprego que geram (Ferreira *et al.*, 2007). A **instabilidade do quadro técnico, sobretudo em organizações sem uma**

fonte financeira estável (dependentes de receitas de projetos, por exemplo)¹³³, constitui, no entender de representantes de OTS e de organizações federativas, **um forte entrave à fixação de competências e de consolidação das práticas de avaliação:**

A questão dos financiamentos temporários dos projetos também prejudica em termos de organização, porque muitas vezes as pessoas, deixando de haver financiamento, saem, a equipa nem sempre é a mesma e é preciso tempo para aprenderem (...). E na verdade os quadros não são sempre os mesmos, e tem que se voltar a aprender. Há 3 ou 4 pessoas que estão há mais tempo e se vão mantendo através dos projetos, mas são apenas uma parte da organização. Depois quando as pessoas começam a aprender é que vão embora, porque não temos condições para assegurar a permanência das pessoas. (Entrevista com dirigente de OTS – saúde)

A fragilidade e volatilidade das estruturas de recursos humanos é bastante e não favorece a consolidação das práticas de avaliação nas organizações. (Entrevista com representante de organização de cúpula – 43)

Dada a escassez de recurso humanos, a polivalência e a flexibilidade imperam neste tipo de organizações, com o pessoal remunerado a acumular uma série de funções e a dar resposta a múltiplas solicitações (Ferreira *et al.*, 2007). Nestas condições, a cultura da ação e o desempenho funcional de curto prazo sobrepõe-se ao trabalho de reflexão e análise sobre a intervenção.

5.1.2. Recursos financeiros

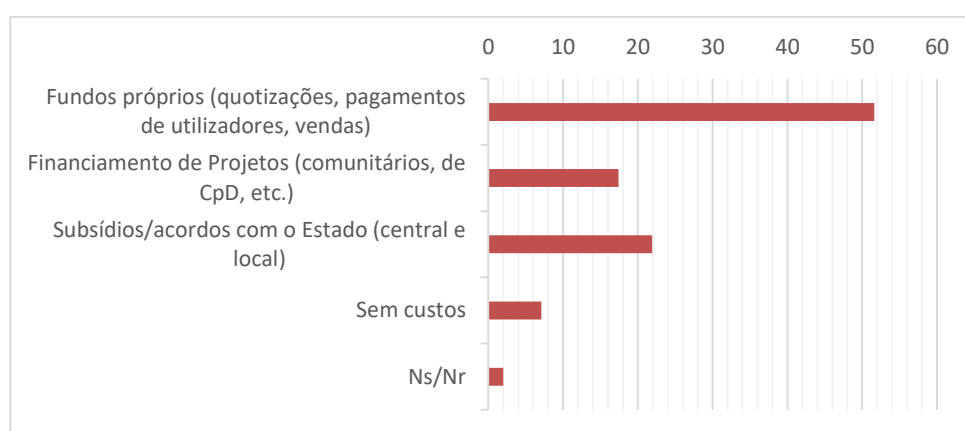
Além de **não especializada**, a atividade de avaliação é **insuficientemente financiada**. A falta de recursos financeiros constitui o maior obstáculo à condução das avaliações nas organizações inquiridas: além de perto de dois terços das organizações nomear a falta de recursos financeiros como um obstáculo importante à avaliação, mais de metade destaca a falta de apoio financeiro por parte das entidades financiadoras/tutelares.

A maioria das entidades inquiridas e entrevistadas refere que **nenhuma das entidades financiadoras suporta financeiramente as suas atividades de avaliação**. Quando questionadas sobre como custeiam a avaliação das suas intervenções (vide gráfico 8), mais de metade das entidades inquiridas afirma contar com fundos operacionais

¹³³ Muitas destas organizações são condicionados pela abertura de concursos a programas de financiamento, que têm um âmbito temporal incerto e uma regulamentação inconstante. Os acordos de cooperação com a Segurança Social para o desenvolvimento de respostas sociais, de que podem beneficiar as IPSS, contribuem, apesar de tudo, para a “estabilidade das receitas das organizações”, uma vez que a Segurança Social comparticipa mensalmente a instituição, constituindo uma fonte financeira estável, em combinação com o financiamento privado das famílias (Martín *et al.*, 2014: 100-101).

próprios (quotizações, participações de utentes, vendas): “Isso vem do nosso budget normal, a não ser que sejam projetos financiados; senão tem que sair do bolo da casa”. Quando os esforços associados à monitorização e avaliação são suportados pelas entidades financiadoras, na maioria dos casos não se trata de um financiamento específico para esse fim, como ilustra este testemunho: “não temos qualquer financiamento específico para avaliação, o apoio é feito por utente, mas de alguma forma podemos afetar alguns destes recursos à avaliação.”

Gráfico 9. Principal fonte de financiamento da avaliação (%)



Fonte: Inquérito por questionário.

De qualquer das formas, nas situações em que as organizações se socorrem de fundos “externos” para fazer face às despesas com as atividades de monitorização e avaliação fazem-no sobretudo no âmbito de acordos com o Estado Central ou Local (programas, subsídios, protocolos/acordos de cooperação) ou no âmbito de projetos específicos – que correspondem às principais fontes de financiamento global das organizações. Embora na maioria dos casos, sobretudo quando se trata de subsídios ou acordos com o Estado, não exista um montante de financiamento destinado especificamente à avaliação, algumas linhas de financiamento, em particular associadas ao desenvolvimento de projetos, existe afetação de verbas especificamente para esse fim. A este nível sobressaem no estudo alguns programas de financiamento específicos, designadamente, os projetos de Cooperação/Educação para o Desenvolvimento, financiados pelo Instituto Camões, nos quais existe uma rubrica específica do financiamento consagrada ao acompanhamento e avaliação. O mesmo acontecia com os projetos

financiados no âmbito da iniciativa comunitária EQUAL, como vimos atrás. Além da avaliação ser obrigatória, o financiamento previa verbas específicas para a avaliação (interna ou externa). No caso dos projetos POPH, não havendo uma rubrica específica para avaliação (embora se tratasse de uma despesa elegível)¹³⁴, nem existindo da parte do programa uma estratégia proactiva no sentido de incentivar a realização de exercícios de avaliação estruturados (o que é ilustrado designadamente pela não obrigatoriedade de apresentação de relatório de avaliação), muitas entidades sequer “arriscam” solicitar verbas específicas para a atividade de avaliação aquando da candidatura. Isto mesmo indicam os resultados do inquérito: perto de dois terços das entidades (62%) com projetos POPH (num total de 253 entidades) afirmam não ter pedido financiamento específico para as atividades de avaliação¹³⁵. Nos casos em que essa verba foi solicitada, foi frequente a sua não aprovação, como também admite responsável do POPH:

Quem tem uma certa dimensão acaba por reservar nos seus custos gerais uma verba específica. Agora, quem pediu exclusivamente ou especificamente um montante para esse efeito, é provável que não tenha sido financiado. Por essa razão, porque não havia uma natureza de despesas elegíveis especificamente para esse efeito. (...) (Entrevista com responsável dos órgãos de gestão POPH)

Não surpreende que as entidades assumam uma atitude passiva e se limitem ao que é “obrigatório”, dado que não têm a avaliação no topo das suas preocupações, não têm esta prática internalizada, nem as entidades financiadoras assinalam claramente a relevância dos exercícios de avaliação das intervenções financiadas. Os testemunhos que se seguem são paradigmáticos:

Acabamos por andar aqui a dar resposta àquilo que são as exigências de quem financia. Em termos de avaliação, se os programas não são muito exigentes, se não nos financiam, nós não temos capacidade para fazer um investimento tão grande a esse nível porque acabamos por não ter retorno. (Entrevista com dirigente de OTS – 20)

Normalmente diz-se que a avaliação é muito importante, e eu também acho, mas tem que se consumir tempo de determinada atividade e da instituição. Mas se nós não conseguimos ter recursos para ter as coisas a funcionar não podemos dar prioridade à avaliação. Em determinados projetos o financiamento é tão

¹³⁴ Apesar de ser uma despesa elegível, ela estava subsumida nos outros encargos gerais do projeto, não havendo, como no passado (no antecessor POEFDS existia uma rubrica específica para avaliação), uma rubrica específica para o efeito. No entender da gestão do programa, a presidir a opção de não discriminação da avaliação na estrutura de despesas elegíveis esteve a intenção de simplificar os procedimentos e não de desincentivar a avaliação.

¹³⁵ Das 90 entidades (36%) que solicitaram verbas para avaliação, mais de 80% obteve aprovação.

escasso que temos que dar prioridade à avaliação e depois, se houver dinheiro, vem a avaliação. (Entrevista com técnico de ONG – 14).

Em geral, os resultados dos estudos conduzidos noutros países levam a conclusões semelhantes, ainda que com algumas variações significativas em termos de valores, dependendo do país ou dos termos em que é colocada a questão do financiamento da avaliação. Tal como nas OTS portuguesas, nas OTS italianas é muito pouco frequente a disponibilização de financiamento específico para as atividades de avaliação por parte das entidades financiadoras/tutelares (Melloni, 2014). Já nas OTS inglesas e norte-americanas, a generalidade dos estudos (Ellis e Gregory, 2008; Innovation Network, 2010, 2012, 2016; Ógáin, Lumley e Pritchard, 2012) parece apontar para um maior suporte financeiro às atividades de monitorização e avaliação das intervenções apoiadas, ainda que ele seja considerando desproporcionado quando comparado com o que se exige às organizações em termos de avaliação¹³⁶.

5.2. Ambiente interno: recetividade e suporte da liderança

Análises anteriores têm chamado a atenção para a importância do papel da direção (estatutária ou executiva) na capacitação em avaliação. O estudo de Hoole e Patterson (2008) revela que apesar de os esforços das entidades financiadoras em criar capacitação serem importantes, o compromisso da liderança organizacional é decisivo na transformação do papel da avaliação de um relato e prestação de contas básico num processo contínuo de aprendizagem organizacional. Esta transformação reivindica, segundo o estudo, o comprometimento da liderança com o desenvolvimento de uma cultura de avaliação e aprendizagem e de uma infraestrutura que a suporte. O papel da

¹³⁶ Segundo o inquérito às OTS inglesas conduzido por Ógáin, Lumley, e Pritchard (2012), cerca de um terço das organizações afirma que pelo menos parte dos seus financiadores incorpora fundos específicos para monitorização e avaliação. Já outro inquérito conduzido no Reino Unido (Ellis e Gregory, 2008) no qual a questão relativa ao financiamento das atividades de avaliação é colocada de forma distinta (não remetendo necessariamente para apoios especificamente vocacionados para atividades de avaliação) a percentagem associada ao suporte financeiro das entidades financiadoras é mais expressiva, abrangendo cerca de dois terços das OTS inquiridas. Inquéritos similares aplicados nos EUA pela Innovation Network (2010, 2012, 2016) revelam resultados similares, apontando para a inexistência de qualquer apoio das entidades financiadoras para conduzir exercícios de avaliação em mais de um terço das OTS. Mais pessimistas são os resultados do inquérito a uma amostra de OTS norte-americanas conduzido por Carman e Fredericks (2008): quando questionadas sobre a principal fonte de financiamento da avaliação, quase dois terços das organizações identificaram fundos internos/operacionais e quase nenhuma (menos de 1%) recebeu financiamento específico para o efeito.

liderança das organizações e as prioridades organizacionais estabelecidas são também considerados nos estudos da Innovation Network (2010, 2012, 2016). A falta de apoio da liderança das organizações é apontada, nos três inquéritos¹³⁷, por duas em cada cinco organizações como barreira que inibe a capacidade para integrar processos de avaliação, sendo considerado um desafio significativo a enfrentar por perto de um terço das entidades inquiridas.

Embora no inquérito conduzido no âmbito deste estudo a falta de apoio da direção para conduzir exercícios de avaliação seja indicado enquanto obstáculo importante por um volume menos significativo de OTS (perto de 15%)¹³⁸, a **importância do papel da direção no impulsionamento e sustentação das práticas de avaliação** sobressai nas entrevistas, em particular com profissionais de avaliação, organizações representativas do setor e com representantes não dirigentes de organizações de base. O envolvimento da direção surge amiúde identificado por estes grupos de entrevistados como um fator crítico para o reforço da capacitação em avaliação e para o efetivo impacto destes exercícios nas organizações.

Há muitas circunstâncias nas organizações em que os líderes são um pouco ausentes e servem um pouco de travão na própria dinâmica da organização. No geral, circunstâncias em que os líderes estavam comprometidos e eram facilitadores, a mudança nas organizações era mais fluída, e noutras circunstâncias pouco ou nada aconteceu. (...) (Entrevista com profissional de avaliação – 39)

Uma liderança aberta, por exemplo, é uma dimensão facilitadora de um processo de avaliação. Uma liderança fechada, no sentido de querer ter todo o protagonismo ou até todo o controle de tudo o que se passa na organização é, regra geral, uma administração ou gestão que faz uma avaliação de conveniência. (Entrevista com representante de organização de cúpula – 45)

Além disso, embora se encontre entre os obstáculos menos aludidos no inquérito, a falta de suporte da liderança é, a seguir à falta de conhecimentos e competências em avaliação, o constrangimento que apresenta a correlação estatística mais forte com a

¹³⁷ Os estudos têm na base o inquérito às OTS norte-americanas realizado em dois momentos: 2009 e 2011.

¹³⁸ Um dado importante para ler estes resultados respeita ao cargo que o/a respondente ocupa na organização: oito em cada dez ocupam posições de liderança – 44% são pessoas da direção estatutária e 37% ocupam lugares de coordenação/direção técnica. Ora, é razoável esperar que assumam uma postura menos crítica na análise do seu papel e mais dificilmente se assumam como elemento obstaculizador das práticas de avaliação. De facto, o teste Qui-quadrado de Pearson releva diferenças estatisticamente significativas ($p=0,003$) de posicionamento relativamente ao apoio da liderança consoante o lugar ocupado na organização pela pessoa respondente. O apoio insuficiente da liderança é considerado um “obstáculo importante” por 9% de representantes da direção estatutária e por 21% do pessoal sem funções de chefia e como “obstáculo menor” por 26% dos primeiros e 39% dos segundos.

intensidade/abrangência¹³⁹ da avaliação nas organizações. Ou seja, **a recetividade da liderança é mais diferenciadora da adesão à avaliação do que fatores como a disponibilidade de recursos financeiros ou do acesso a apoio de *stakeholders* externos.**

De resto, nas entrevistas o maior **fechamento e antiguidade das lideranças** (pouca rotatividade dos corpos dirigentes) é apontado como (potencialmente) obstaculizador do investimento nestas práticas, que reivindicam novas formas de fazer e de estar nas organizações e uma efetiva abertura à crítica e à mudança. Os casos que se trazem de seguida à colação refletem estes bloqueios:

Algumas pessoas responsáveis pela avaliação não estão preparadas totalmente para esse exercício. (...) Os dirigentes estão cá há muitos anos e isso também tem impactos. Havia potencial para haver uma maior apropriação dos resultados em termos de gestão. (...) O facto de a direção serem pessoas de uma certa idade e algumas delas com debilidades até de saúde... (...) (Entrevista com direção técnica de OTS – 13)

Depois, o que teve um impacto grande e que foi quase uma guerra iraquiana, foi a Escuta e Envolvimento de *Stakeholders*. É muito difícil a pessoa numa organização como a nossa que tinha a mesma direção [há muito tempo]... (...) Isso traz desgaste e uma certa forma de viver as coisas que, depois, a pessoa se sente pouco confortável a colocar a sua imagem nas mãos e à mercê do pensamento de todas as pessoas. (Entrevista com dirigente de OTS – 8)

O tipo de liderança autocrática, caracterizada pela dependência muito acentuada do/a presidente/diretor/a, é apresentada nas entrevistas como um entrave à introdução de uma cultura de avaliação, ao constranger a autonomia das organizações, entendida enquanto capacidade de afirmação autónoma da instituição na prossecução dos seus objetivos relativamente aos indivíduos que a constituem, em particular as lideranças (Ferreira *et al.*, 2007: 133):

Há muitas ONG portuguesas que vivem pela iniciativa de uma determinada pessoa (...), que normalmente é o presidente ou o diretor executivo. Se essa pessoa não tiver como prioridade a questão da avaliação será mais difícil que a

¹³⁹ Foi pedido às entidades que seleccionassem uma das seguintes opções, que descrevem intensidade das práticas de monitorização e avaliação: 0 – Não fazemos qualquer avaliação das nossas atividades; 1 – Fazemos muito pouca monitorização e avaliação das nossas atividades; 2 – Fazemos alguma monitorização e avaliação das nossas atividades; 3 – Monitorizamos e avaliamos a generalidade das nossas atividades; 4 – Monitorizamos e avaliamos todas as nossas atividades.

Para avaliar se a abrangência/abrangência da avaliação se correlaciona com cada um dos 8 constrangimentos listados (vide gráfico 8) recorreu-se ao coeficiente de correlação de Spearman (variáveis medidas numa escala ordinal) (vide tabela 9, anexo 5).

organização incorpore práticas de avaliação mais permanentes na própria ONG.
(Entrevista com representante de organização de cúpula – 43)

A dimensão e o grau de profissionalização das organizações, estando vinculadas à capacidade de desenvolvimento do trabalho técnico de forma qualificada e autónoma, potenciará a autonomia organizacional em relação às lideranças e à capacidade da organização responder às condições atuais de complexificação das relações entre estas organizações e os agentes do Estado e do mercado (Ferreira *et al.*, 2007: 133).

Lá está, a profissionalização leva a que a maior parte dessas organizações já não dependa dessa pessoa que as fundou, da sua boa vontade e visão estratégica para a questão da avaliação. Mas há ainda em muitos casos em que isso é uma limitação, nomeadamente nas ONGD mais pequenas, a falta de visão estratégica das pessoas que lideram a própria organização. (Entrevista com representante de organização de cúpula – 43)

De facto, nas organizações entrevistadas em que se verifica o envolvimento em processos de gestão e avaliação ambiciosos em termos de investimento e impacto na organização é comum a referência ao apoio e envolvimento das lideranças nos processos:

Temos tido sempre o apoio incondicional da direção. Mesmo quando foi apresentado o projeto para requerer a certificação EQUASS, foi apoiado e validado como uma mais-valia para a instituição. (Entrevista com coordenação técnica de OTS – 19)

Além do suporte da liderança das organizações, uma cultura organizacional favorável à condução de exercícios de avaliação beneficia de uma **atitude positiva das pessoas envolvidas** nos processos de recolha e análise da informação (Preskil e Boyle, 2008), característica de uma **cultura de aprendizagem** (Preskil e Torres, 1999). Apesar de a perceção da avaliação como uma ameaça se encontrar, a par com a falta de suporte da liderança, entre os obstáculos menos prementes para as organizações, ele é identificado por quatro em cada dez (43%) como obstáculo à avaliação. Nas entrevistas este constrangimento perde também centralidade em favor dos obstáculos associados aos recursos e ao ambiente externo. Todavia, quando nas entrevistas se suscita que as entidades discorram sobre a recetividade da avaliação, surgem frequentemente referências a resistências, sobretudo iniciais, motivadas por fatores diversos: pela forte **carga administrativa** associada à formalização destes processos, por dificuldades de **gestão do tempo** e de recursos entre uma multiplicidade de atribuições, pela perceção da avaliação como processo **punitivo**, pela perceção da avaliação como um algo **obrigatório/imposto, automático**, ou pela visão da avaliação como **desperdício de**

recursos, ampliada pela percepção da **inutilidade destes exercícios**. Os excertos que se seguem, colhidos de entrevistas com *stakeholders* diversos, são reveladores de algumas destas percepções:

Nós vivemos muito a gestão do próprio processo, tanto da associação como dos projetos, e depois sem fazer uma gestão de tempo para fazer esta avaliação. Às vezes o tempo não é muito. E depois a própria preocupação é pela sobrevivência, tanto da execução como da associação. (Entrevista com dirigente de IPSS – 25)

Não é líquido para toda a gente que a avaliação é importante, porque isto da avaliação leva muito tempo. (...) Se calhar porque também não veem os resultados dessa avaliação plasmados nas intervenções futuras. (...) É uma postura mais pragmática: “ninguém liga, para que é que eu agora hei-de estar a perder tanto tempo?”; “isto é uma chatice, a avaliação dá trabalho”. (Entrevista com coordenação técnica de ADL – 26)

Há ainda muito uma visão de inutilidade que ainda é muito generalizada: isto não serve para nada. (...) É mais um elemento da carga administrativa. (Entrevista com representante do observatório do QREN - 1)

Há uma cultura instaurada de que de facto é importante as pessoas irem ter connosco para avaliar como é que as práticas são feitas, mas que as coisas são um bocado inconsequentes. Há sempre esses dois extremos. As pessoas reconhecem que é importante, mas acreditam vai ser inconsequentes, que não há coisas concretas e objetivas que vão decorrer daquilo. (Entrevista com dirigente de ONG – 28)

A relutância no comprometimento com a avaliação, quando motivada pela percepção da sua **inutilidade**, parece ser mais frequente quando enquadrada em processos de **prestação de contas ascendente** (a entidades tutelares/financiadoras) focados na execução física e financeira de projetos, isto é, centrada em realizações/*outputs* e de carácter descritivo (é o caso dos relatórios de execução dos projetos), na linha do previsto na primeira alínea da segunda hipótese do estudo. A compreensão das mais-valias e vantagens associadas às práticas de avaliação é, pois, decisiva para mitigar ou ultrapassar reações adversas, como se ilustra neste excerto:

A 1ª reação é sempre de algum acréscimo de trabalho de secretária. Mas depois, lá está, pelo menos quem acompanha os projetos de perto, sente depois que, nos projetos e candidaturas que se apresentam no futuro, a mais-valia de ser implementado numa fase anterior todo o processo de avaliação. Porque existem dados, existem evidências que servem depois para a justificação dos projetos seguintes, e implementação de forma mais eficaz no desenvolvimento das atividades futuras. (Entrevista com direção técnica de OTS – 19)

De resto, a garantia de que o investimento nas práticas de monitorização e avaliação tem retorno (externo) é um pressuposto importante para que a avaliação seja bem acolhida e sustentada:

Mas também porque isso, na perspetiva de muitas delas, não lhes traz grandes benefícios ao nível do crescimento do seu trabalho e de conseguir mais financiamentos por apostarem numa avaliação estratégica maior. Há casos de organizações portuguesas que fizeram relatórios que tinham uma componente muito forte de avaliação e gestão estratégica. E fizeram isso durante 1, 2 anos, com custos muito elevados, como não viram retorno ou qualquer benefício acrescido de continuar a fazer esse trabalho desistiram. (Entrevista com representante de organização de cúpula – 43)

A perceção da avaliação como uma **ameaça**, por seu turno, associa-se particularmente à prestação de contas interna aos órgãos dirigentes/sociais, sobretudo quando ela implica decisões sobre recursos humanos¹⁴⁰. Apesar de a avaliação de desempenho individual não constituir o foco do estudo, nas entrevistas ela é frequentemente trazida para ilustrar a visão da avaliação enquanto **processo punitivo** e por isso particularmente sensível nas organizações, como fica patente no testemunho colhido numa entrevista com organização de base do TS:

É rara a pessoa que sente a avaliação de desempenho como um momento para poder melhorar em vez de ser um castigo ou punitiva. (...) Aquilo era praticamente uma avaliação que apenas era para melhorar. (...) Nem havia castigos e muito menos havia incentivos. (...) Mas as pessoas sentiam sempre um momento muito desgastante, de comparação entre os vários colegas e gerou sempre imenso conflito. (Entrevista com pessoal dirigente de OTS – 8)

Curiosamente, vemos que dentre o conjunto de práticas de gestão, controlo e avaliação listadas no inquérito¹⁴¹, a elaboração de **relatórios** de contas e de planos anuais de atividades encontra-se entre as práticas mais associadas a perceção da avaliação como uma ameaça por parte dos membros da organização, eventualmente por se tratar de um requisito legal, podendo ser encarada como uma imposição. Associa-se também em especial a processos de **auditoria externa** para verificação da conformidade com normas e

¹⁴⁰ Para avaliar se a perceção da avaliação como uma ameaça varia de acordo com os usos dos resultados das avaliações recorreu-se ao Teste do Qui-quadrado de Pearson. A análise estatística permite afirmar que aquela perceção varia de forma estatisticamente significativa consoante os usos da avaliação (*p-value* = 0,025). Vemos que o reporte aos órgãos sociais, as decisões sobre recursos humanos e o não uso da avaliação se destacam na maior consideração da perceção da avaliação como ameaça por parte dos membros da organização (respetivamente, 46%, 50% e 69%).

¹⁴¹ Para avaliar se a perceção da avaliação como uma ameaça varia de acordo com os processos de gestão e avaliação conduzidos regularmente pelas organizações, recorreu-se ao Teste do Qui-quadrado de Pearson. A análise estatística permite afirmar que aquela perceção varia de forma estatisticamente significativa consoante os processos listados (*p-value* = 0,000). Vemos que a elaboração de relatórios anuais, as auditorias técnicas da iniciativa das entidades financiadoras/tutelares e as análises SROI se destacam na maior consideração da perceção da avaliação como ameaça por parte dos membros da organização (respetivamente, 44%, 40% e 50%).

procedimentos. É o caso das auditorias técnicas (onde se incluem as visitas de acompanhamento da Segurança Social) da iniciativa das entidades tutelares/financiadoras, particularmente “stressantes” para as organizações por representarem uma ameaça à continuidade do financiamento ou mesmo da atividade.

Estes resultados deixam emergir uma **ambivalência** na forma como a avaliação é acolhida nas organizações: embora a sua boa receptividade esteja dependente da perceção de que se trata de um investimento consequente, a antevisão de potenciais punições é desmobilizadora para as pessoas que participam nestes exercícios.

5.3. O ambiente externo: iniciativas de suporte e capacitação

Especialistas e profissionais de avaliação entrevistados são unânimes na constatação da **debilidade da oferta** – e procura – **de iniciativas de capacitação em avaliação dirigidas a organizações do terceiro setor**, apesar de se reconhecer o crescimento destas iniciativas:

Eu diria que em Portugal evoluímos bastante do que era há 10 anos o mercado de avaliação, mas estamos muito longe... (...) Em 2000, ia ter dificuldade em pagar um euro que fosse para eu dar formação sobre avaliação. Se se avançar 10 anos atrás, ainda menos. Hoje em dia, eu digo-lhe que as pessoas me procuram 50% por questões de planeamento e 50% avaliação, mas não era esta a realidade, temos vindo a melhorar. (Entrevista com profissional de avaliação – 37)

No mesmo sentido ressoam as vozes de dentro das organizações do setor, especialmente de pessoal técnico envolvido na condução de exercícios de avaliação aprofundados e centrados nos efeitos das intervenções, fora do enquadramento dos sistemas de gestão da qualidade e do planeamento anual ou estratégicos. Algumas pessoas auscultadas em entrevista lamentam a indisponibilidade de **recursos técnicos** no domínio da avaliação (plataformas de partilha de metodologias e instrumentos, guias, manuais, ferramentas, etc.) que apoiam as organizações neste tipo de exercício:

O problema é que isto ainda é um trabalho exploratório. Não há manuais que apoiem esse trabalho a nível nacional, faltam indicadores que ajudassem a fazer este retrato pelas diferentes áreas. Do ponto de vista do esquema existe algum

– ainda que pouco - apoio, mas do ponto de vista do conteúdo não há nada. Ao nível da educação, saúde, ... (Entrevista com técnica de IPSS – 4)

O carácter relativamente pouco frequente de iniciativas de carácter formativo¹⁴²

no domínio da avaliação (de nível organizacional e/ou de iniciativas/projetos específicos) dirigidas às OTS é reiterado por pesquisa exploratória conduzida *online* no âmbito deste estudo (vide capítulo metodológico). Neste levantamento apenas foram considerados eventos especificamente vocacionados para a formação em avaliação (nas suas diversas feições e modalidades). Ficam por isso de fora desta análise cursos de formação em gestão organizacional ou em gestão da intervenção social mais abrangentes, por exemplo, ações de formação-ação para a qualificação das entidades da economia social¹⁴³, ou cursos de formação/pós-graduação em gestão e capacitação organizacional que, podendo integrar conteúdos sobre avaliação, não incidem especificamente sobre a temática.

Os conceitos e as práticas associados à avaliação do desempenho e do impacto das OTS estão, porém, cada vez mais presentes no espaço público, sob tradições teóricas e profissionais diversas, chegando a um público cada vez maior e mais diversificado. No total, entre 2010 e 2016, foram identificadas **51 iniciativas de formação**, relativas a eventos de duração variável subordinados a objetivos de aprendizagem acerca da temática da avaliação/medição do desempenho nas organizações do terceiro setor. Perto de um quarto (22%) destes eventos são de **curta duração** (inferior a 8 horas), abrangendo um conjunto diverso de formatos, desde “conferências” e “palestras”, a “seminários” e “workshops”, em geral de entrada livre e com a presença de oradores/as convidados/as¹⁴⁴.

O grosso das iniciativas formativas identificadas (70%) corresponde a **eventos de média duração** (com duração entre 12 e 24 horas), organizados segundo formatos geralmente designados por “ação de formação” e, menos frequentemente, “workshop”,

¹⁴² Iniciativas formativas com ou sem ciclo de estudos ou graduação associados, conferentes ou não conferentes de grau. Foram abrangidas na pesquisa iniciativas de formação contínua – cursos de formação, *workshops*, conferências/palestras/seminários – e cursos de formação graduada e pós-graduada (cursos de licenciatura, pós-graduação/especialização, mestrado ou doutoramento) ministradas por instituições académicas.

¹⁴³ Por exemplo, iniciativas no âmbito do Programa de Formação-Ação para Entidades da Economia Social, do POPH.

¹⁴⁴ A título ilustrativo destaco a *workshop* “Monitorização e Avaliação de Impacto Social”, que decorreu em Coimbra em 2015, a palestra “Avaliação de Impacto e Desenvolvimento Local” que decorreu no mesmo ano em Lisboa, e as conferências sobre avaliação do impacto social (2013 e 2015) realizadas no âmbito do Programa Impacto Social.

“seminário de especialização” ou “unidade curricular” (no caso de iniciativas dinamizadas por instituições académicas). São dirigidas especialmente a pessoal técnico e dirigente de organizações do terceiro setor e implicam geralmente algum grau de participação financeira, assim como uma certificação de competências¹⁴⁵.

Apesar de a pesquisa ter abrangido iniciativas de **longa duração**, designadamente relativas a cursos de formação graduada e pós-graduada (licenciatura, mestrado, pós-graduação ou doutoramento), a pesquisa apenas referenciou 4 ações formativas específicas sobre avaliação com duração superior a 24 horas.¹⁴⁶ Tal não invalida que o tema da avaliação não integre ações formativas mais abrangentes e de maior duração, como é o caso dos cursos de formação ou pós-graduação em gestão (estratégica/da qualidade) de organizações do terceiro setor. Todavia, como foi referido acima, só foram contabilizadas na análise ações especificamente subordinadas ao tema da avaliação. Em todo o caso, a presença de instituições de ensino superior na oferta de iniciativas formativas especificamente vocacionadas para a avaliação é modesta (cobre 16% do total da oferta). A pouca atenção dada à avaliação no domínio da intervenção social nos currículos e outras iniciativas do mundo académico é particularmente salientada por organizações de cúpula e profissionais de avaliação entrevistados.

Nas faculdades não há tradição nessa área. (...) Se percorrermos os currículos académicos, há muita pouca ênfase dada à questão da avaliação (...) Esta é uma área ainda um bocadinho a descobrir. (...) Mas as universidades têm um papel: de criar os currículos, de investir em técnicos que possam ajudar a dar essas cadeiras. (...) Ensinar às pessoas que isto é relevante, que é um instrumentos de

¹⁴⁵ Estes tipos formativos incluem, por exemplo, o “Seminário de Especialização em Conceção e Avaliação de Projetos” promovido pelo Instituto para as Políticas Públicas e Sociais do Instituto Universitário de Lisboa, a *workshop* “Avaliação de Resultados e Impacto Social” dinamizada pela Fundação Eugénio de Almeida e a ação de formação “Desenvolver competências para a elaboração do Plano de Avaliação das Estratégias de Desenvolvimento Local” promovida pela Federação Portuguesa de Desenvolvimento Local Minha Terra.

¹⁴⁶ Trata-se de unidades curriculares inseridas em planos de estudo de cursos pós-graduados (pós-graduações e mestrados), à exceção de uma unidade curricular de um curso de licenciatura (em Sociologia) – “Gestão e Avaliação de Projectos de Intervenção Social” (da Faculdade de Economia Universidade de Coimbra). Inclui também a unidade “Metodologias de Gestão e Avaliação de Projetos” dos Mestrados em Sociologia e em Empreendedorismo e Inovação Social da Universidade da Beira Interior, a disciplina de “Gestão e Avaliação de Programas e Projectos Sociais” do Mestrado em Intervenção Social, Inovação e Empreendedorismo das Faculdades de Economia e de Psicologia e Ciências da Educação da Universidade de Coimbra, e a unidade curricular “Análise e Avaliação de Políticas Públicas” do Mestrado em Economia Social da Universidade Católica do Porto.

trabalho para irem para o mercado de trabalho, (...) que é diferente de todos. (Entrevista com profissional de avaliação – 37)

Depois eu também acho que até ao nível das relações com o meio académico, podia haver um trabalho mais articulados, através ou de seminários, ou *workshops*, eventos onde estas questões fossem mais trabalhadas. Porque a academia tem o discurso mais legitimado. E ser a academia a dar conta desta necessidades e dos resultados daqui decorrentes, também tem muito peso. (Entrevista com representante de organização de cúpula – 42)

Um resultado relevante e contundente desta análise respeita exatamente ao **tipo de entidade promotora das iniciativas de formação**. Constatou-se que as organizações de cúpula do terceiro setor foram as principais impulsionadoras, cobrindo perto de dois terços oferta formativa (60%). Dentre as diversas organizações representativas do setor, destaca-se claramente a EPAN (Rede Europeia Anti Pobreza), responsável pela dinamização de um terço do total dos eventos formativos, com várias edições regionais/anuais de ações de formação sobretudo vocacionadas para a avaliação no âmbito de projetos de intervenção social¹⁴⁷. As organizações de base do terceiro setor, por seu turno, e muito em especial as fundações¹⁴⁸, são responsáveis pela organização e promoção de cerca de 17% dos eventos. A presença de entidades privadas na promoção de eventos formativos subordinados à temática da avaliação das organizações/intervenções sociais é diminuta (foi identificado apenas um caso) ainda que seja generalizado o envolvimento de técnicos/as de empresas de consultoria na dinamização das ações¹⁴⁹.

No que respeita ao conteúdo das ações, sobressaem as abordagens associadas à **avaliação de projetos**, sendo **menos frequente** a avaliação da organização como um todo. A maioria destes eventos visa reforçar a capacidade de dirigentes e técnicos/as das organizações do terceiro setor para planear e implementar processos de monitorização e avaliação de projetos sociais. Merece também destaque, **sobretudo desde 2013, o enfoque das iniciativas de formação na medição/avaliação do impacto**, muito na esteira dos desafios colocados pelo novo quadro comunitário de fundos estruturais “Portugal 2020”, e da crescente popularidade da filantropia de risco/de impacto no nosso país.

¹⁴⁷ Alguns exemplos de ações de formação: “Avaliação e Monitorização de Projetos de Intervenção Social”, “Avaliação de Resultados e Impactos Sociais”, “Ciclo de Projeto e Avaliação de Resultados: teoria da mudança e metodologia SROI”.

¹⁴⁸ Um terço dos eventos dinamizados por organizações de base do TS foram promovidos pela Fundação Eugénio de Almeida.

¹⁴⁹ A generalidade dos eventos contam com a presença de oradores/as, facilitadores/as ou formadores/as externos às entidades promotoras.

Quando questionadas sobre o seu papel atual na promoção de práticas de avaliação nas organizações que representam, as **organizações de cúpula auscultadas em entrevista referem-se essencialmente à sua atuação no campo da formação**. No topo das referências destas entidades (e das organizações de base do setor) estão as ações de formação-ação enquadradas no Programa de Formação - Ação para Entidades da Economia Social:

Este projeto [de qualificação organizacional] permitiu já formar para a qualificação uma centena de organizações (...) O enfoque que é dado depois na formação é nos processos de avaliação também: avaliação de desempenho, avaliação da satisfação dos clientes, de avaliação de reconhecimento organizacional e por aí a diante. (Entrevista com representante de organização de cúpula – 45)

Apesar de excluídas da análise por não se tratar de ações especificamente vocacionadas para capacitação em avaliação, estas **iniciativas de formação e consultoria** destacam-se nos inquéritos e entrevistas conduzidos no âmbito do estudo como assumindo um papel relevante neste domínio. Com efeito, 1 em cada 10 OTS inquiridas (sobretudo IPSS) destaca a participação nestas iniciativas de formação, orientadas para a melhoria dos processos de gestão e reforço das competências de dirigentes e trabalhadores/as, como particularmente relevantes para a integração de práticas de avaliação. Estes projetos foram promovidos por entidades representativas do terceiro setor e envolveram um número muito significativo de organizações de base do setor (mais de 3000, entre 2007 e 2013, segundo relatórios de execução do POPH). Os projetos Q3 - Qualificar o 3º setor¹⁵⁰, QUAL-IS – Qualificar a Intervenção Social¹⁵¹, PRIO – Promover Redes Inovação e Oportunidades¹⁵² e FAS – Formação Ação Solidária¹⁵³ são os exemplos de iniciativas conduzidas no âmbito daquela tipologia de intervenção mais aludidos pelas pessoas auscultadas. No conjunto destas iniciativas o relevo dado à componente da avaliação é diferenciado, assim como são diversos os enquadramentos e enfoques da avaliação nesses programas de formação, ainda que sobressaia (ao contrário das iniciativas

¹⁵⁰ O Q3 envolve uma parceria de âmbito nacional constituída pela MINHA TERRA (Federação Portuguesa das Associações de Desenvolvimento Local), FENACERCI (Federação Nacional de Cooperativas de Solidariedade Social), CPCCRD (Confederação Portuguesa das Coletividades de Cultura, Recreio e Desporto), Associação In Loco, AEP (Associação Empresarial de Portugal), IEBA (Centro de Iniciativas Empresariais da Beira Aguireira) e UTAD (Universidade de Trás-os-Montes e Alto Douro).

¹⁵¹ Promovido pela EAPN - Rede Europeia Anti-Pobreza/Portugal.

¹⁵² Promovido pela EAPN - Rede Europeia Anti-Pobreza/Portugal.

¹⁵³ Promovida pela CNIS.

de formação específicas em avaliação)¹⁵⁴ a incidência **na organização como um todo**. Encontramos desde o enquadramento da avaliação na gestão da qualidade (enfoque na auditoria, medição de resultados e avaliação da satisfação), na gestão de recursos humanos (enfoque na avaliação do desempenho individual) até ao seu enquadramento no planeamento estratégico (enfoque em sistemas de avaliação do desempenho organizacional – análise SWOT, BSC) e no desenvolvimento de projetos (enfoque na avaliação de projetos).

Apesar de não ter sido identificada na dinamização de nenhuma das iniciativas de formação-ação aludidas, o papel da plataforma das ONGD na capacitação em avaliação das entidades que representa merece também destaque particular. Além das iniciativas de formação específicas em avaliação de projetos com enfoque nos impactos e da valorização da componente da avaliação nas ações de formação sobre planeamento estratégico que dinamiza¹⁵⁵, a plataforma aposta na sensibilização das suas associadas para a importância de planearem e avaliarem o seu trabalho, não apenas numa perspetiva de processos, mas também de resultados e impactos:

Temos promovido vários debates e várias formações, inclusivamente para as nossas associadas, que procuram dar-lhe instrumentos para que possam utilizar e fazer essas avaliações: construir os seus planos estratégicos e avaliá-los. Temos feito formações de planeamento estratégico e formações de avaliação com essa perspetiva do impacto. E não só formações, como seminários, *workshops*, trazendo cá outras organizações internacionais promover isso. (...) Um dos processos que está a ser iniciado agora (...) é a criação de um código de conduta, de ética (...) que todas as nossas associadas têm que cumprir. E isso terá obviamente uma componente de avaliação da organização e do trabalho que fazem, o cumprimento de determinados procedimentos, de determinados valores. (Entrevista com representante de organização de cúpula – 43)

A ideia da centralidade das organizações de cúpula do setor na capacitação das organizações em avaliação sugerida pela sua importância na dinamização de iniciativas formativas no domínio da avaliação é **relativizada pelos resultados do inquérito por**

¹⁵⁴ Além do POPH, e embora com uma extensão mais modesta, este tipo de iniciativas de capacitação de OTS tem também cabimento em um dos quatro domínios de atuação do Programa Cidadania Ativa¹⁵⁴ - Reforço da Eficácia da Ação das ONG (domínio C) –, nomeadamente no âmbito da implementação de “inovações organizacionais e de gestão nas ONG para a melhoria da governação e da qualidade e diversidade dos serviços prestados”. Neste domínio de atuação foram apoiados 31 projetos que contam com o envolvimento direto de 53 ONG¹⁵⁴.

¹⁵⁵ Estas formações incluem dois pontos dedicados à avaliação: (i) Análise do impacto e da eficácia dos vários programas/atividades e (ii) Análise do impacto e da eficácia da organização.

questionário às OTS. Com efeito, o cruzamento da existência/aceso a competências internas e/ou externas em avaliação com a pertença a plataformas nacionais representativas não revela interações estatisticamente significativas¹⁵⁶. Ou seja, em termos gerais, **a maior estruturação do subsector em que atuam e/ou a maior institucionalização do setor/área em que operam, não parece ter impacto significativo em termos da existência de *expertise* em avaliação nas organizações**¹⁵⁷.

A análise desagregada por organização representativa revela, não obstante, algumas *nuances* com significado estatístico¹⁵⁸. Vemos que, coerentemente com o aventado na análise qualitativa, as entidades associadas **da plataforma das ONGD têm maior acesso a competências, quer internamente, quer externamente**. Nos antípodas estão, como seria de esperar, as entidades associadas da Confederação Portuguesa das Coletividades de Cultura, Recreio e Desporto, com reduzido acesso a competências em avaliação. As entidades associadas da ANIMAR e FENACERCI acompanham a Plataforma das ONGD no acesso a competências externas. No caso das entidades pertencentes à FENACERCI, as competências externas estarão sobretudo associadas aos processos de certificação da qualidade, a que estas entidades se associam particularmente (80% estão envolvidas em processos de certificação numa base regular, face ao valor global de 26%). No enquadramento destes processos, que as entidades identificam amiúde com a avaliação

¹⁵⁶ Para avaliar se as competências em avaliação são influenciadas pela participação da organização em projetos recorreu-se ao Teste do Qui-quadrado de Pearson, considerando-se uma probabilidade de erro de tipo I de 0,05. A análise estatística permite afirmar que quer as competências internas, quer as externas não variam de forma estatisticamente significativa consoante a existência de projetos (respetivamente, $\chi^2_{kw} = 2,692$, $p\text{-value} = 0,101$ e $\chi^2_{kw} = 0,368$, $p\text{-value} = 0,544$).

¹⁵⁷ A ausência de uma correlação positiva entre competências em avaliação e pertença a organizações de cúpula não será indiferente ao peso das coletividades (que apresentam valores de adesão a práticas de avaliação relativamente reduzidos) na amostra de entidades que pertencem a organizações de cúpula: constituem perto de um terço (31%) da amostra, e são as entidades mais enquadradas neste tipo de plataforma (72%), seguidas das IPSS (67%).

¹⁵⁸ Para avaliar se as competências em avaliação (internas e externas) são influenciadas pela pertença às diversas plataformas recorreu-se ao Teste do Qui-quadrado de Pearson para cada uma das plataformas, individualmente, considerando-se uma probabilidade de erro de tipo I de 0,05. Ao nível das competências internas, verifica-se a existência de diferenças estatisticamente significativas entre entidades que pertencem e não pertencem à CPCCRD e à Plataforma das ONGD (respetivamente, $\chi^2_{kw} = 23,503$, $p\text{-value} = 0,000$ e $\chi^2_{kw} = 4,886$, $p\text{-value} = 0,027$). Em relação ao acesso a competências externas, verificam-se diferenças estatisticamente significativas entre entidades que pertencem e não pertencem à ANIMAR, FENACERCI, Plataforma das ONGS e CPCCRD (respetivamente, $\chi^2_{kw} = 5,113$, $p\text{-value} = 0,024$; $\chi^2_{kw} = 7,773$, $p\text{-value} = 0,005$; $\chi^2_{kw} = 4,886$, $p\text{-value} = 0,027$; e $\chi^2_{kw} = 13,202$, $p\text{-value} = 0,000$).

do desempenho organizacional, é comum recorrer-se a consultoria externa para apoiar nas designadas “auditorias internas” da qualidade. A própria plataforma pretende disponibilizar às suas associadas este tipo de recursos, como nos foi reportado por representante:

O outro papel é sermos nós próprios auditores internos. Estamos aqui a preparar recursos para fazer auditoria interna às nossas próprias associadas; não auditoria externa que isso há outras entidades que o fazem. Depois, é o papel formativo. Nós temos um núcleo para a formação e qualificação, e uma das linhas onde fazemos formação é no domínio da avaliação.

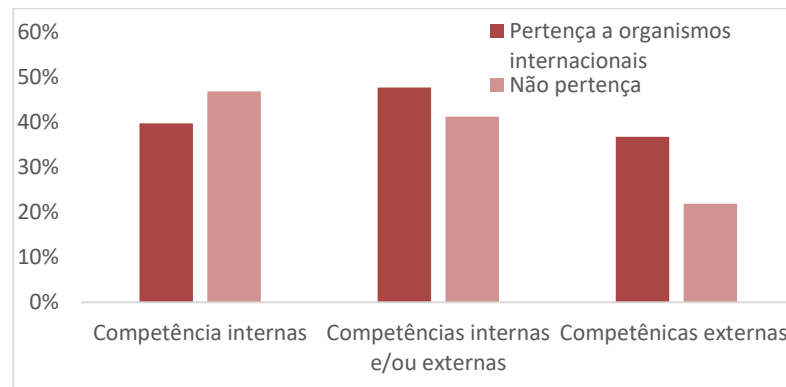
Já no caso das entidades associadas da plataforma das ONGD e da ANIMAR, o acesso a competências externas dá-se muito por via do desenvolvimento de projetos, nomeadamente para o apoio à elaboração de relatórios e/ou de avaliações formais de projeto.

Se a pertença a plataformas nacionais tem um impacto reduzido sobre as competências avaliativas, a pertença a organismos/redes internacionais parece exercer uma influência mais significativa (vide gráfico abaixo), especificamente no recurso a profissionais externos. Com efeito, as organizações que afirmam pertencer a organismos ou redes internacionais tendem a ter um acesso significativamente maior a competências externas no domínio da avaliação¹⁵⁹. A mobilização de competências externas parece ser, nestes casos, sobretudo colocada ao serviço da avaliação conduzida no âmbito dos projetos.¹⁶⁰

¹⁵⁹ O teste qui-quadrado de Pearson permite-nos concluir que, para uma probabilidade de erro de 5%, o recurso a profissionais externos varia de forma estatisticamente significativa consoante a pertença a organismos/redes internacionais ($\chi^2_{kw} = 8,910$, $p\text{-value} = 0,003$).

¹⁶⁰ Dentre o conjunto de práticas de gestão, controlo e avaliação listadas no inquérito, as relativas aos projetos (especificamente, elaboração de relatórios e avaliações formais de projetos e conceção de modelos lógicos) encontram-se entre as que mais separam organizações pertencentes e não pertencentes a redes/organismos internacionais.

Gráfico 10. Competências em avaliação segundo pertença a plataformas/redes internacionais



Fonte: Inquérito por questionário.

As entrevistas com organizações de base e de cúpula sustentam a ideia de que estas redes tendem a ter um impacto importante nas organizações em termos de capacitação em alguns domínios específicos, por efeitos de “isomorfismo institucional” (Ferreira *et al.*, 2007), sobretudo de tipo mimético e tipo normativo. O testemunho que se segue é emblemático:

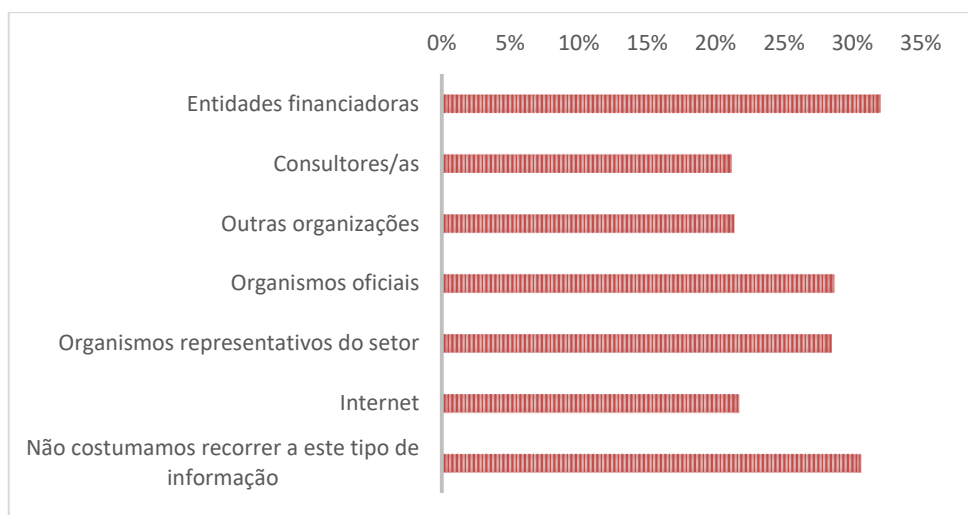
O 1º plano estratégico da X foi aprovado em 1993, numa altura em que se falava muito pouco de PE, mesmo ao nível empresarial aqui em Portugal, começava a falar-se. E esta ideia do PE surge por causa da federação da nossa organização no [organismo internacional], que é uma ONG das maiores mundiais e tem procedimentos de gestão em geral muito avançados. Eles têm uma tradição anglo-saxónica, e portanto foram eles... (Entrevista com dirigente de ONG – 9)

A ideia da relevância das organizações de cúpula do setor para a capacitação das organizações em avaliação sugerida pela sua importância na dinamização de iniciativas formativas no domínio da avaliação é também colocada em perspetiva quando questionamos as organizações sobre as principais fontes de informação sobre avaliação a que costumam recorrer (vide gráfico 10).

Nas respostas a esta questão sobressai desde logo o peso significativo (perto de um terço) das entidades que afirma **não recorrer a este tipo de recurso**, o que por si já é indicador da fraca relevância da avaliação num conjunto importante de OTS. Depois, na linha do proposto na segunda hipótese do estudo, a **importância das entidades financiadoras e dos organismos oficiais**, destacados globalmente por perto de metade das organizações, é **reveladora do peso que estas entidades têm no enquadramento e definição das abordagens de avaliação seguidas pelas organizações de base do terceiro**

setor, em especial quando se limitam a responder aos requisitos do financiamento no que toca à avaliação. Em muitos casos, a informação sobre avaliação disponibilizada nos documentos orientadores dos programas (regulamentos, diretrizes para a apresentação de candidaturas, manuais de procedimentos, guias de utilização, formulários, modelos, etc.), constitui o principal ou único recurso para conduzir exercícios de avaliação nas organizações. Em consequência, a medida da profundidade das avaliações conduzidas surge associada à do investimento das entidades financiadoras/tutelares em recursos técnicos e instrumentos de apoio à condução de avaliações por parte das organizações tuteladas/financiadas.

Gráfico 11. Fontes de informação sobre avaliação



Fonte: Inquérito por questionário.

Observamos que, enquanto **algumas entidades se limitam a definir**, de forma mais ou menos vaga, **requisitos de avaliação** nos regulamentos ou avisos de abertura de candidaturas, sem produção de documentos orientadores ou indicação de abordagens, mecanismos ou metodologias de avaliação possíveis, outras, com maior ou menor intensidade, **facultam algum tipo de apoio à condução dos exercícios de avaliação**, seja através da divulgação e/ou produção de documentos orientadores/de apoio (guias, guiões, manuais, modelos de relatório, etc.), seja através de sessões de esclarecimento, sensibilização ou formação sobre avaliação.

Os dois programas estudados em profundidade neste trabalho – PIC EQUAL e POPH – enquadram estes **dois tipos de posicionamento** no que toca às iniciativas de

capacitação em avaliação das entidades beneficiárias por parte das entidades financiadoras. O apoio técnico disponibilizado às entidades beneficiárias para avaliação dos projetos do POPH foi reduzido, particularmente quando comparado com o investimento feito pelo PIC EQUAL, embora mais uma vez se ressalve a natureza distinta de ambos os programas – o POPH é um programa *mainstream* enquanto o PIC EQUAL era um programa experimental. Não foram produzidos, no âmbito do POPH, documentos de apoio/orientadores da atividade de avaliação nos projetos, nem tiveram lugar ações de formação/esclarecimento/sensibilização dirigidos a promotores dos projetos para apoiar as atividades de avaliação. Apenas tiveram lugar sessões de informação e esclarecimento sobre projetos abrangidos pelas tipologias de intervenção. No âmbito do QREN são disponibilizados alguns recursos e informação de suporte (nomeadamente o Guia para a Avaliação do Desenvolvimento Socioeconómico – EVALSED, da Comissão Europeia), mas são sobretudo orientados para a avaliação dos próprios POs e menos especificamente à avaliação dos projetos. Já no caso do EQUAL, a Gestão do Programa promoveu uma grande diversidade de iniciativas específicas de apoio à avaliação dos projetos: produziu e disponibilizou referenciais e instrumentos para a avaliação (guias para a autoavaliação¹⁶¹, para os balanços de competências¹⁶², para a avaliação independente¹⁶³, para a validação de produtos¹⁶⁴ e para a avaliação da disseminação¹⁶⁵); organizou sessões de informação/formação prática a promotores/avaliadores; disponibilizou recursos financeiros e técnicos para a realização das avaliações nos projetos; produziu sínteses sobre os exercícios de avaliação realizados pelos projetos (autoavaliação, balanços de competências) e organizou sessões de “devolução” de resultados; elaborou e propôs aos projetos um sistema de informação e monitorização das ações¹⁶⁶ (Figueirôa, 2010).

¹⁶¹ “Guia da Auto-avaliação dos Projectos” (EQUAL, 2003) e “Guia da auto-avaliação dos projectos – 2ª Fase” (EQUAL, 2005a).

¹⁶² Guia do “Balanço de Competências” (Estêvão, 2002) e Guia do “Balanço de Competências – 2ª Fase” (Estêvão, 2005).

¹⁶³ “Avaliação independente dos projectos EQUAL” (EQUAL, 2005b).

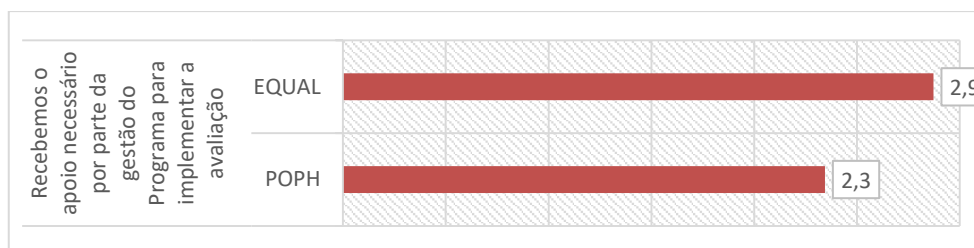
¹⁶⁴ “Produtos EQUAL: Validar para disseminar” (EQUAL, 2004).

¹⁶⁵ “Validação de Produtos Inovadores” (EQUAL, 2007a) e “Disseminar Produtos: Referencial de Apoio” (EQUAL, 2007b).

¹⁶⁶ Construído enquanto referencial de base ou comum a todos os projetos, envolvendo as componentes da execução financeira e física.

O posicionamento das entidades inquiridas relativamente ao apoio recebido para implementar a avaliação por parte da gestão de cada um dos programas é revelador do maior investimento do EQUAL face ao POPH¹⁶⁷.

Gráfico 12. Perceções sobre suporte recebido para implementação da avaliação (POPH versus EQUAL)*



Fonte: Inquérito por questionário.

Nota: *Valores médios: 1 = Discordo totalmente; 2 = Discordo; 3 = Concordo; 4 = Concordo totalmente

No entender da generalidade dos atores relevantes no domínio da avaliação entrevistados no âmbito do estudo – pessoas envolvidas na gestão de programas FSE, especialistas e profissionais na área da avaliação e representantes de OTS e de organizações de cúpula do setor – **nunca como no PIC EQUAL se foi tão longe nas exigências colocadas a projetos de intervenção social no domínio da avaliação, produzindo orientações e instrumentos de apoio, promovendo a formação, encorajando e financiando o investimento na avaliação.**

Tal entendimento aponta para a corroboração da alínea a) da terceira hipótese do estudo. De resto, a afirmação da avaliação como componente essencial do desenho e implementação dos projetos financiados está plasmada quer nos documentos enquadramentos do programa quer nos guias e outros documentos de apoio aos utilizadores do programa.

Não obstante, em alguns programas mais recentes encontramos um esforço considerável no sentido de **providenciar as orientações básicas para produzir os exercícios de monitorização e/ou avaliação requeridos.** É o caso dos Programas CLDS/CLDS + (Contratos Locais de Desenvolvimento Social), no âmbito dos quais são facultados

¹⁶⁷ Para testar se existiam diferenças entre POPH e EQUAL nos valores médios de adesão à afirmação foi realizado o teste de Wilcoxon para uma mediana. Concluímos, com uma probabilidade de erro de 5%, que os valores médios de adesão a ambas as afirmações são significativamente diferentes no POPH e no EQUAL ($p=0,002$).

formulários de monitorização e um guião para a autoavaliação dos projetos para apoiar a recolha e análise da informação sobre os resultados das ações. É também o caso dos **Projetos de Cooperação e Educação para o Desenvolvimento de ONGD**, financiados pelo Instituto Camões, em que as Diretrizes para a Apresentação de Propostas são acompanhadas de documentos de apoio úteis ao planeamento da atividade de avaliação dos projetos, concretamente, documentos orientadores do trabalho de recolha dos indicadores e de elaboração de quadro lógico. Do mesmo modo, o **Programa Cidadania Ativa**, gerido pela Fundação Calouste Gulbenkian, providencia algum suporte à condução das atividades de monitorização e avaliação por parte dos promotores por via das diretrizes contidas no manual do promotor que é disponibilizado (vários documentos auxiliares de preenchimento de formulários e relatórios contêm indicações sobre o tipo de informação de monitorização/avaliação a recolher). O **Programa Escolhas**, por seu turno, no seu *handbook* “Fazer Escola com o Escolhas”, dedica alguns artigos à questão da avaliação, produzindo orientações específicas sobre a metodologia a seguir na “avaliação por objetivos” preconizada pelo programa, além de propor um conjunto de indicadores para objetivos de resultado em cada uma das áreas de intervenção. Em nenhum destes programas encontramos, todavia, uma mobilização em torno da capacitação das entidades beneficiárias para a avaliação das suas intervenções comparável à verificada no contexto do PIC EQUAL, onde a avaliação foi uma área fortemente investida com referenciais conceptuais e operativos.

Em todo o caso, os resultados do estudo empírico¹⁶⁸ são reveladores da **influência da participação em projetos na existência de ou no acesso a competências de avaliação nas organizações, especialmente internas**. Enquanto perto de dois terços das entidades inquiridas (64%) com projetos financiados pelo FSE afirma contar com competências (internas ou externas) específicas em avaliação, pouco mais de um terço das organizações (38%) sem projetos afirma ter acesso a este tipo de competências.

¹⁶⁸ Para avaliar se as competências em avaliação são influenciadas pela participação da organização em projetos recorreu-se ao Teste do Qui-quadrado de Pearson. A análise estatística permite afirmar que quer as competências internas, quer as externas variam de forma estatisticamente significativa consoante a existência de projetos (respetivamente, $\chi^2_{kw} = 21,822$, $p\text{-value} = 0,000$ e $\chi^2_{kw} = 27,351$, $p\text{-value} = 0,000$).

Fora do contexto dos projetos – e no enquadramento dos **sistemas de gestão de qualidade** – é de relevar o contributo para a capacitação das organizações, também no domínio da medição/avaliação do desempenho dos serviços prestados, dos instrumentos técnicos produzidos e disponibilizados pelo Instituto da Segurança Social – no âmbito do Programa de Cooperação para o Desenvolvimento da Qualidade e Segurança das Respostas Sociais.¹⁶⁹ Os **Referenciais de Gestão da Qualidade das Respostas Sociais**, que definem os requisitos para a avaliação da qualidade, apoiam o desenvolvimento de ferramentas para a medição/avaliação do desempenho das respostas sociais, sobretudo ao nível da avaliação do grau de satisfação de clientes, de colaboradores/as e de entidades parceiras dos serviços e equipamentos de apoio social. Além de modelos de questionários de avaliação da satisfação de diferentes *stakeholders* para as diversas respostas sociais, é proposto uma série de indicadores úteis para medir o desempenho em cada um dos “manuais de processos-chave” de prestação de serviços que são disponibilizados. Ainda que não isentos de críticas¹⁷⁰, **o potencial destes referenciais para apoiar as OTS nos seus objetivos de análise do desempenho não deverá ser menosprezado**, particularmente ao nível dos **processos operacionais**. Muitas são as organizações (IPSS) envolvidas no estudo que afirmam recorrer a estes referenciais não apenas para organizar internamente as valências, em termos de uniformização de processos, mas também para definir objetivos e indicadores de qualidade para as respostas sociais, para encetar processos de auscultação da satisfação dos *stakeholders*, ou ainda para produzir relatórios de avaliação de planos individuais de beneficiários/as.

Por último, há que destacar a proliferação, nos últimos anos, de **iniciativas de apoio ao desenvolvimento da agenda de impacto social em Portugal**, de que o Programa Impacto Social é exemplo. Promovido pela Cooperativa António Sérgio para a Economia Social, pela Fundação Montepio e pela consultora 4Change, encontra-se na sua terceira edição (foram envolvidas 20 organizações nas duas primeiras edições), e tem como objetivo capacitar as entidades da economia social para a avaliação de impacto das suas

¹⁶⁹ Os projetos de formação e consultoria promovidos por organizações de cúpula do setor no âmbito do POPH – a que nos referimos acima – potenciaram a implementação deste programa, capacitando as entidades para a aplicação dos referenciais.

¹⁷⁰ Veja-se, por exemplo Lucas *et.al* (2012).

ações através da metodologia SROI - *Social Return on Investment*. As entidades selecionadas recebem formação introdutória e mentoria para desenvolver a análise SROI da organização ou projeto específico. Na última edição do Programa, foi introduzida uma primeira etapa de seleção, que prevê a participação em seminários *online* (“webinars”) e a resposta a um conjunto de desafios com vista à seleção das organizações que participarão na etapa de mentoria para a demonstração do potencial de impacto dos seus projetos. Uma das organizações envolvidas neste programa que foi auscultada no âmbito deste trabalho, atesta a importância da participação nesta iniciativa no reforço das competências para realizar exercícios de avaliação:

Eu já tento ver onde posso ir buscar indicadores disto ou daquilo, de que forma é que posso escrever o projeto em que pelo meio eu tenho uma série de indicadores avaliativos que me possam no final, além do inquéritos e entrevistas, ou do *feedback* que possa ter, o que é que eu posso acrescentar mais ali. As coisas já são construídas com outro olhar. O que a formação de SROI nos trouxe foi este olhar. (Entrevista com dirigente de ADL – 29)

Na esteira do Programa Impacto Social, a empresa nele envolvida – a 4Change – criou uma plataforma *online* – *samia*¹⁷¹ – onde disponibiliza uma base de dados com indicadores e “aproximações” financeiras em diversos domínios temáticos, e que poderá ser um instrumento relevante para quem pretende aceder rapidamente a informação operativa para mensuração do impacto, em particular através da metodologia SROI.

Apesar da relevância do conjunto de iniciativas alvitradas, ela não tem sido suficiente para colmatar a escassez de *expertise* em avaliação.

Outros estudos têm corroborado estas conclusões (e.g., Franco, 2015), sendo a falta de competências em avaliação frequentemente enquadrada na falta de competências de gestão (estratégica e/ou operacional) da direção e/ou do corpo técnico.

Mas faltam competências às pessoas da gestão social, faltam-nos competências para uma gestão global. E a gestão tinha que ser feita por um gestor. (...) Estas organizações do terceiro setor efetivamente ainda não estão habituadas a ser de gestão, porque são sempre da área social. (Entrevista com dirigente de IPSS – 8) Depois, o corpo técnico, o diretor técnico, a quem são cometidas todas as responsabilidades de gestão, têm depois muito défice de formação na área da gestão, têm formações muito mais na área social e muito menos na gestão. (Entrevista com representante de organização de cúpula – 44)

¹⁷¹ *Stakeholders*, atividades, mudanças, indicadores e aproximações financeiras: SAMIA. Acesso em 6 de Fevereiro de 2016, disponível em <http://samia.pt/>.

De resto, este enquadramento evidencia mais uma vez a **aproximação ao discurso gestor sobre a avaliação** e o afastamento das origens da avaliação enquanto ciência social e competências que se lhes associam.

5.4. Capacidade para fazer avaliação: retratos organizacionais

Entre o conjunto de variáveis que em diversos estudos são apontadas como influentes da capacidade avaliativa encontram-se a dimensão financeira, antiguidade organizacional, perfil da atividade e fonte de financiamento (Ellis, 2007; Carman e Fredericks, 2008, 2009, 2010; Hoefler, 2009; Verbruggen, Christiaens e Milis, 2011; Eckerd e Moulton, 2011). As respostas ao inquérito às OTS permitem-nos estabelecer algumas tendências no que toca a essa influência.

Uma questão central colocada às organizações (através de inquérito por questionário e entrevista) a partir da qual é possível revelar padrões relativos à **capacidade para fazer avaliação**, respeita aos **constrangimentos/obstáculos à condução dos exercícios de avaliação**. A questão colocada no inquérito permitiu recolher informação sobre diversas dimensões associadas a esta capacidade, concretamente: **recursos humanos** (“Tempo limitado da equipa de trabalho” e “Escassez de competências em avaliação”); **recursos organizacionais** (“Recursos financeiros insuficientes” e “Problemas com a recolha de informação”); e **contexto organizacional interno** (“Apoio insuficiente por parte da liderança” e “Perceção da avaliação como uma ameaça”); e **externo** (“Falta de apoio técnico por parte das entidades financiadoras/tutelares” e “Falta de apoio financeiro por parte das entidades financiadoras/tutelares”).

No sentido de resumir a informação presente nas variáveis associadas aos obstáculos à avaliação em componentes principais e, neste caso, sobretudo para explorar as suas correlações, utilizei a Análise Categórica de Componentes Principais (ACCP)¹⁷².

¹⁷² A Análise de Componentes Principais (ACP) é um dos métodos estatísticos mais usados quando se pretende analisar dados multivariados, permitindo transformar um conjunto de variáveis originais, intercorrelacionadas, num novo conjunto de variáveis não correlacionadas, as componentes principais. O objetivo mais imediato da ACP é verificar se existe um pequeno número das primeiras componentes principais que seja responsável por explicar uma proporção elevada da variação total associada ao conjunto original. Não sendo o objetivo deste método explicar a covariância/correlação entre as variáveis, ele permite constatar e explorar estas relações, já que as componentes principais são combinações lineares ponderadas

Foi possível reduzir a informação em quatro componentes ortogonais que explicam 70% da variância total das variáveis originais (vide tabela 11). A primeira componente tem o maior peso explicativo e é determinada pela generalidade dos constrangimentos listados. Uma vez que quase todas as dificuldades elencadas se correlacionam fortemente com esta dimensão, podemos dizer que ela define “**constrangimentos múltiplos**”. Na segunda componente, que designo de falta de “**suporte financeiro e técnico**”, destacam-se, exatamente, os *scores* elevados da falta de apoio técnico e financeiro por parte das entidades financiadoras. A terceira componente, por seu turno, associa à **falta de recursos financeiros a falta de tempo da equipa**. Por fim, a quarta componente, define a **falta de conhecimentos e competências em avaliação**, a que se associa a **falta de apoio por parte da liderança**.

Tabela 11. Obstáculos à avaliação (Componentes extraídas da ACCP)

	Componente			
	Múltiplos	Suporte financeiro e técnico	Recursos financ. e humanos	Competências e liderança
Recursos financeiros insuficientes	,508	,436	,493	,169
Tempo limitado da equipa de trabalho	,405	-,385	,755	-,173
Escassez de competências em aval.	,612	-,436	,136	,438
Apoio insuficiente da liderança	,574	-,171	-,412	,436
Falta de apoio técnico por parte das entidades financiadoras/tutelares	,671	,449	-,206	-,150
Falta de apoio financeiro por parte das entidades financiadoras/tutelares	,610	,628	-,027	-,162
Problemas com a recolha de info.	,461	-,430	-,172	-,621
Perceção da aval. como uma ameaça	,566	-,330	-,301	-,084
<i>Eigenvalue</i>	2,481	1,448	1,164	,882
Variância explicada	29,0%	16,9%	13,6%	10,3%

Fonte: Inquérito por questionário.

As características da estrutura organizacional, à exceção do perfil da atividade e dimensão organizacional, têm uma influência estatisticamente pouco significativa nas diversas tipologias de constrangimentos extraídas na análise de componentes principais¹⁷³.

das variáveis originais. Neste caso em concreto pretende-se sobretudo analisar as correlações entre as variáveis originais.

¹⁷³ Com o objetivo de perceber que características da estrutura organizacional se associam significativamente a cada uma das componentes, recorreu-se à Anova a um fator (one-way), mediante a

Similarmente ao observado com a análise agregada, a análise individualizada de cada um dos obstáculos revela o **fraco poder explicativo das características da estrutura organizacional**, ainda que todos os constrangimentos, à exceção do tempo limitado da equipa de trabalho, sejam influenciados por pelo menos um dos fatores inseridos na análise (apenas a antiguidade não exerce influência sobre qualquer dos constrangimentos).

Apresenta-se, de seguida, um quadro que sintetiza aquelas análises, com base apenas nos resultados estatisticamente significativos ($p\text{-value} < 0,05$). As características da estrutura organizacional são apresentadas por ordem decrescente de influência na adesão aos diversos obstáculos.

Quadro 16. Obstáculos à avaliação: perfis de OTS

	Maior adesão	Menor adesão
Múltiplos	Pequenas/Médias	Grandes Filantropia
Escassez de recursos financeiros e humanos	Coletividades → ONG Pequenas Expressão	IPSS/ADL/Outras Grandes → Médias Serviços
Escassez de competências técnicas	Coletividades/IPSS	ONG
Falta de suporte da liderança organizacional	Receitas públicas → Próprias	Filantropia
Perceção da avaliação como ameaça	IPSS/Outras Serviços	ONG → ADL Mistas

Fonte: Inquérito por questionário.

Desde logo, é possível notar que, em geral, cada tipo de constrangimento é influenciado por fatores organizacionais distintos, evidenciando **a natureza distinta dos diversos domínios de capacitação**. Depois, globalmente, as organizações de **grande dimensão financeira são as mais imunes** à coexistência de múltiplas fragilidades associadas à capacidade avaliativa. As organizações de grande dimensão, com maior robustez financeira, têm mais recursos para investir em processos que suportam a ação direta, sendo que a avaliação pode ser muito consumidora de tempo e recursos financeiros

utilização dos scores associados às componentes extraídas na ACCP. Uma vez que a análise agregada dos obstáculos à avaliação se revelou limitada no propósito de criar padrões distintivos de organizações, analisou-se também de forma desagregada cada um dos obstáculos elencados, através do teste não paramétrico de Kruskal-Wallis, seguido da comparação múltipla da média das ordens.

(designadamente quando enquadrada em iniciativas de gestão abrangentes, como o planeamento estratégico ou a gestão da qualidade). Além disso, como notam Hwang e Powell (2009: 285), o planeamento e a avaliação poderão constituir veículos de coordenação e controlo organizacional, uma vez que as entidades de maior dimensão tendem a envolver num leque mais extenso e variado de atividades e quem as dirige tende a estar mais envolvido na administração e menos na ação direta.

Dentre os diversos perfis de organização, as coletividades são as que enfrentam as maiores dificuldades, seja em termos de **recursos financeiros e de tempo** (a par com as ONG), seja em termos de competências técnicas. São exatamente as **coletividades** e as **ONG** – entidades em que as funções de expressão são mais marcadas¹⁷⁴, com menor dimensão financeira e com maior peso do voluntariado (Ferreira *et al.*, 2007) – que maiores dificuldades enfrentam na **dimensão organizacional dos recursos para avaliação**. As coletividades, em particular, além de uma gestão pouco profissionalizada, encontram-se também muito dependentes de fundos próprios¹⁷⁵ para conduzir exercícios de monitorização/avaliação.

As **IPSS** partilham com as **coletividades** a centralidade da **escassez das competências** no conjunto dos constrangimentos – constrangimento que afeta menos as ONG (particularmente dos setores da participação cívica e defesa de causas e da cooperação internacional). Trata-se de entidades com características genericamente muito distintas, quer em termos de perfil de funções e atividades, quer em termos de dimensão e perfil de financiamento. Enquanto as coletividades são caracterizadas pela forte componente expressiva das suas atividades (sobretudo culturais, desportivas e recreativas), as IPSS estão mais associadas a um propósito mais instrumental de prestação de serviços. Enquanto as IPSS se destacam entre as organizações com maior dimensão quer em termos de volume de receitas quer em termos de número de trabalhadores/as, as coletividades são tendencialmente de pequena dimensão e assentes no trabalho

¹⁷⁴ A grande maioria das ONG (mais de 80%) e todas as coletividades inquiridas desenvolvem atividade em domínios marcadamente de expressão (de expressão de valores e interesses profissionais, culturais, políticos ou sociais e defesa de causas e direitos). 28% das ONG e 83% das coletividades não desenvolve atividades de prestação de serviços.

¹⁷⁵ Os fundos próprios são a principal fonte de financiamento das atividades de avaliação para 83% das coletividades.

voluntário. Enquanto o financiamento público é proeminente nas IPSS, as coletividades estão sobretudo dependentes de receitas próprias. Já no que toca à capacidade técnica parecem existir alguns constrangimentos análogos que poderão fomentar a aproximação destes dois tipos de entidade no domínio das competências em avaliação. Embora contanto globalmente com um maior número de colaboradores/as, o peso de trabalhadores/as com qualificações superiores no total do pessoal ao serviço é limitado quando comparado com outro tipo de entidades, nomeadamente ONG e ADL (Ferreira et al., 2007: 135). Esta observação encontra-se em linha com a perceção de uma pessoa entrevistada envolvida na avaliação de um programa de formação dirigido a OTS que realça as fragilidades técnicas destes tipos de entidades:

Aquilo que nós temos observado é que as IPSS, misericórdias, CERCIS, etc., são normalmente organizações com um corpo de colaboradores bastante variáveis mas normalmente bastante grande (...), com um corpo qualificado e profissional relativamente reduzido, com exceção das CERCIS, que já têm mais assistentes sociais, psicólogos... Depois, temos as ADL que são normalmente relativamente pequenas em nº de colaboradores, mas um corpo de colaboradores muito profissionalizado e qualificado. (...) E depois temos as coletividades que funcionam muito à base de voluntários, pouco profissionalizadas. (...) São pequenas associações, [assentes em] trabalho voluntário, (...) têm menor massa crítica de colaboradores... a apetência não é tão elevada. (Entrevista com profissional de avaliação – 39)

Além da questão dos obstáculos à avaliação, outras questões colocadas no inquérito por questionário¹⁷⁶ permitem-nos acrescentar linhas ao perfil das organizações nesta dimensão essencial da capacitação em avaliação: a **expertise em avaliação**. A partir de um conjunto de análises estatísticas (bivariadas e multivariadas)¹⁷⁷, e apenas com base nos resultados estatisticamente significativos ($p\text{-value} < 0,05$), foi possível traçar perfis de organizações associadas positiva ou negativamente à disponibilidade (interna e/ou externa) de competências no domínio da avaliação. A negrito são destacadas as associações estatisticamente significativas na análise multivariada (vide quadro 17).

¹⁷⁶ As questões do inquérito mobilizadas para esta análise foram: “Em geral, quem são os agentes faz atividades de avaliação” e “A sua organização conta, internamente, com pessoas(s) com competências específicas em avaliação?”

¹⁷⁷ Enquanto a análise bivariada (com recurso ao coeficiente de correlação Phi) das características organizacionais influentes das competências de avaliação – dimensão financeira, antiguidade, principal fonte de receitas, tipo de entidade, perfil/âmbito de atividade – permite-nos conhecer a efeito desses fatores por si só, individualmente considerados, a análise multivariada (regressão binominal – 3 modelos de regressão para cada um dos tipos de competências) permite-nos considerar todos os fatores em simultâneo, dando-nos uma perspetiva dinâmica de interinfluência.

Quadro 17. Competências internas e externas em avaliação: perfis de OTS

	DIMENSÃO	TIPO ENTIDADE	RECEITAS	FUNÇÕES	IDADE
Competências internas	Grandes	ONG/IPSS	Apoio público (projetos)	Serviços Serviços Sociais/Educação Desenvolvim. e Ambiente	Antig. intermédia
Competências externas	Grandes/ Médias	ONG/ADL/ IPSS	Apoio público	Mistas → Serviços Atividades Internacionais/ Desenvolvim. e Ambiente	Antig. intermédia
Avaliação profissional (competências externas e/ou internas)	Grandes/ Médias	ONG/ADL/ IPSS	Apoio público (projetos)	Serviços/ Mistas Participação Cívica e Defesa de Causas/ Serviços Sociais/Educação /Desenvolvim. e Ambiente	Antig. intermédia

Fonte: Inquérito por questionário.

Vemos que a dinamização de exercícios de avaliação por pessoas com competências específicas em avaliação é tendencialmente mais frequente em organizações de **maior dimensão**, com **funções de serviço ou mistas** (serviço e expressão) e de constituição **recente** ou **intermédia**, com financiamento predominantemente público (especialmente de **projetos**), sobretudo ONG (do domínio da participação cívica e defesa de causas), mas também ADL e IPSS. As **pequenas** organizações, de constituição mais **antiga**, com funções de **expressão** e **autossustentadas**, em especial as coletividades, encontram-se particularmente vinculadas a processos de avaliação **não especializados**.

Existem, não obstante, **nuances** a este nível, consoante nos reportemos a **competências internas** ou **externas** à organização. Vemos, desde logo, que as ONG se destacam das outras tipologias de OTS na detenção de competências internas, a que se juntam as ADL no recurso a competências externas. Ou seja, são as ONG que estão nas melhores condições para responder aos desafios da monitorização e avaliação em termos de acessibilidade de competências, externas e internas. De resto, são ONG as entidades que menos identificam a escassez de competências em avaliação como obstáculo à condução de exercícios de avaliação, ao contrário das coletividades e IPSS, particularmente

expostas a este constrangimento. Depois, os padrões de antiguidade das entidades que recorrem a competências interna e externamente também são mais convergentes, apontando para a “fragilidade” das entidades de constituição mais antiga, que contam globalmente com menos competências em avaliação. Nos antípodas encontramos as entidades de constituição intermédia, com maior acesso a competências internas e externas. As organizações mais jovens, por seu turno, embora aparentemente motivadas para a avaliação, estão menos capacitadas para aceder a especialistas. Estas tendências são coerentes com os resultados do estudo sobre as influências da profissionalização nas OTS conduzido por Hwang e Powell (2009), segundo o qual as organizações de constituição mais recente são mais “racionalizadas” que as de constituição antiga, estando mais orientadas para a contratação de pessoal especializado nas áreas de suporte ao desenvolvimento das atividades, nomeadamente de gestão, e mais rápidas na adoção de práticas racionalizadas. Por outro lado, por se tratar de organizações menos robustas financeiramente (particularmente as mais jovens)¹⁷⁸, terão menos recursos disponíveis para investir na avaliação (Innovation Network, 2010, 2012, 2016; Eckerd e Moulton, 2011), designadamente para a contratação de pessoal ou para o recurso a consultoria especializada.

O perfil das entidades no que toca ao contexto organizacional interno, especificamente em relação a elementos da **cultura organizacional**, é distinto do associado aos **recursos humanos e organizacionais**. Com efeito, o perfil das atividades desempenhadas e do financiamento são as características organizacionais mais relevantes para compreender aquela dimensão. Vemos que os problemas de **recetividade da avaliação** são particularmente sentidos em organizações prestadoras de serviços (sobretudo IPSS) e menos frequentes em entidades híbridas (sobretudo ONG). As abordagens de monitorização e avaliação privilegiadas pelas IPSS e outras entidades que concentram a sua atividade na prestação de serviços (sociais, de saúde, educativos ou de desenvolvimento) terão influência nesta distribuição. Estas entidades estão especialmente vinculadas a processos de avaliação do desempenho individual que, como vimos atrás, são

¹⁷⁸ Perto de dois terços (61%) das entidades mais jovens da amostra (até 20 anos) concentra-se no escalão de receitas mais baixo (até 99.999 euros/ano).

amiúde invocados pelas organizações para ilustrar a resistência das pessoas envolvidas. Já o exercício de diferentes funções – o fornecimento de serviços e ao mesmo tempo de atividades de expressão de valores, interesses e crenças culturais, espirituais, profissionais ou políticas – parece fomentar a visão da avaliação enquanto experiência positiva, prevenindo posições defensivas da parte dos membros das organizações.

A falta de **suporte da liderança** (outro elemento da cultura organizacional implicado na capacidade avaliativa), por seu turno, é influenciado pelo perfil de receitas das organizações, sendo menos frequente em entidades dependentes da filantropia individual e/ou empresarial. Além de serem organizações especialmente jovens (mais de metade tem até 20 anos), estão particularmente associadas a exercícios de avaliação formais, seja à escala dos projetos, seja à escala organizacional, e à mobilização de ferramentas de avaliação específicas (modelos lógicos e análises SROI). Isto parece dar força à tese da importância do apoio da liderança para conduzir processos de avaliação, sobretudo quando são mais exigentes em tempo e recursos técnicos, como é o caso das práticas alicerçadas na investigação em ciências sociais.

5.5. Síntese: a debilidade dos contextos, recursos e estruturas de avaliação

As abordagens e modelos de avaliação prevaletentes nas organizações têm enquadramento num contexto organizacional débil em termos de estruturas e recursos para fazer avaliação. Desde logo, são raros os casos em que as organizações contam com uma equipa (ou pessoa), departamento ou estrutura interna especificamente dedicada à avaliação. Além de a maioria das organizações não ter capacidade financeira para integrar recursos humanos a trabalhar exclusivamente nesta área, também não parece existir a convicção da necessidade e/ou vantagem de ter pessoas a trabalhar exclusivamente nestas funções. Além disso, a instabilidade do quadro técnico, sobretudo em organizações sem uma fonte financeira estável, constitui um forte entrave à fixação de competências e de consolidação das práticas de avaliação. Os processos de avaliação são, pois, geralmente conduzidos por pessoal interno envolvido na implementação das atividades, sem competências específicas na área. A carência de *expertise* em avaliação é escassamente contrariada pelo ambiente institucional das organizações, o que é revelado pela debilidade

da oferta de iniciativas de caráter formativo no domínio da avaliação dirigidas às OTS – muita embora ela se encontre em crescimento –, e pelo papel limitado que têm desempenhado quer entidades financiadoras quer estruturas representativas do setor em termos de suporte técnico à avaliação, não obstante o peso que as primeiras têm tido na definição das abordagens de avaliação prosseguidas pelas organizações, e a relevância de algumas iniciativas de capacitação promovidas por ambos os tipos de atores.

Além de não especializada, a atividade de avaliação é insuficientemente financiada. A falta de recursos financeiros constitui, de resto, o obstáculo à condução das avaliações mais invocado nas organizações estudadas. Todavia, sendo um constrangimento transversal à generalidade das organizações, é limitado no intuito de estabelecer tendências marcantes da capacidade avaliativa, fazendo emergir outras variáveis mais diferenciadoras das organizações. Do mesmo modo, se robustez financeira é relevante para garantir recursos para avaliação (as organizações maior dimensão tendem a ter mais recursos humanos e financeiros, maior acesso a competências técnicas, e mais infraestruturas de suporte à avaliação), ela deixa de ser diferenciadora do ambiente interno das organizações, nomeadamente no que respeita ao suporte da liderança e a receptividade dos membros à avaliação, evidenciando-se a natureza distinta dos diversos domínios de capacitação. O estudo mostra que os constrangimentos com maior associação aos níveis de adesão à avaliação mais baixos não se colocam tanto ao nível dos recursos - financeiros, materiais ou técnicos -, mas mais ao nível de cultura organizacional. O contexto interno nos níveis de adesão mais modestos à avaliação tenderá a ser marcado por uma cultura organizacional hostil à avaliação, assente numa liderança pouco aberta e comprometida com processos de reflexão, análise e aprendizagem contínua, e/ou numa atitude geral reativa à avaliação por parte dos membros da avaliação (perceção da avaliação como uma ameaça ou como uma perda de tempo). Ao contrário, uma capacidade avaliativa forte será tributária de um clima de abertura, confiança e encorajamento suportado no apoio e/ou envolvimento das lideranças e, por outro lado, potenciada pelo equilíbrio entre os papéis de aprendizagem e de prestação de contas associados à avaliação e pela perceção de utilidade e/ou retorno externo destes exercícios. Na linha destes resultados, o estudo faz emergir uma ambivalência na forma como a avaliação é acolhida nas organizações: se, por

um lado, a perspectiva de sanções associada à condução de exercícios de avaliação abre campo a fenómenos de manipulação ou distorção da informação avaliativa e se associa a visões negativas da avaliação, por outro, a percepção de que a dedicação à avaliação é inútil/não tem retorno é claramente dissuasora deste tipo de investimento, elevando o seu carácter simbólico e cerimonial. Esta ambivalência remete-nos para a consideração das forças externas na própria configuração do contexto organizacional interno, designadamente no que respeita à cultura organizacional.

PARTE III: CONCLUSÕES

Capítulo 6. Perfis e modelos de avaliação nas organizações do terceiro setor

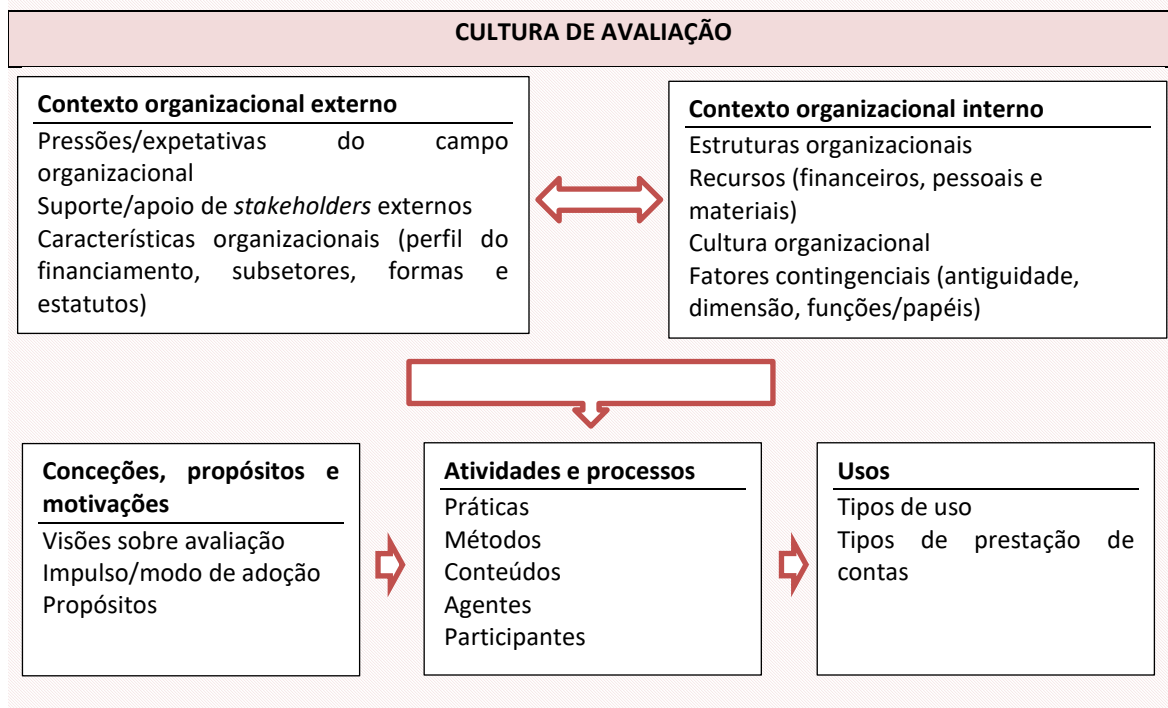
Neste capítulo procura-se fazer uma leitura de conjunto dos resultados explorados na segunda parte do trabalho, mediante a definição de perfis e modelos de avaliação nas organizações do terceiro setor. Tal análise, simultaneamente simplificadora e agregadora, servirá de suporte a uma síntese das informações mais parcelares sobre contextos, práticas e representações sobre avaliação nas OTS e suas implicações em termos de internalização da cultura avaliativa.

6.1. Perfis de avaliação nas organizações do terceiro setor

Partindo das tendências verificadas nos diversos domínios e dimensões associados à adoção de práticas de avaliação nas organizações do terceiro setor nacionais, procede-se agora a uma análise globalizante capaz de – através de procedimentos estatísticos – definir perfis de avaliação.

A definição de perfis foi conduzida com base num vasto conjunto de indicadores relativos a práticas e percepções sobre monitorização e avaliação, numa série de dimensões e sub-dimensões que informam a capacidade avaliativa das organizações e que compõem o modelo analítico de referência que se mobilizou ao longo da investigação. Na figura 7 sintetizam-se as dimensões e sub-dimensões associadas à cultura avaliativa, que sustentaram a definição de perfis organizacionais.

Figura 7. Dimensões e subdimensões associadas à cultura avaliativa nas OTS



Fonte: Elaboração da autora.

As duas primeiras dimensões remetem para a influência do contexto organizacional nas práticas de avaliação. A primeira respeita ao ambiente externo e foca nos requisitos e expectativas de prestação de contas dos diversos *stakeholders*, particularmente das entidades financiadoras/tutelares, assim como na existência ou não de incentivos ou mecanismos de suporte e apoio técnico ou financeiro às organizações. Nesta dimensão são ainda explorados os efeitos das características organizacionais particularmente associadas ao contexto em que operam (perfil do financiamento e formas ou estatutos legais), assim como outras variáveis tradicionalmente associadas às influências miméticas e normativas do ambiente institucional (por exemplo, a pertença a plataformas/redes de organizações, ou posse de uma acreditação/certificação).

A segunda dimensão associa-se ao ambiente interno e é composta por quatro subdimensões. A primeira centra-se nas estruturas e meios pelos quais a avaliação se torna institucionalizada na organização (departamento ou estrutura interna própria dedicada à avaliação, diretrizes e regras comuns, definição de funções e responsabilidades associados à avaliação...); a segunda nos recursos financeiros, pessoais e materiais disponíveis para a avaliação (orçamento para avaliação, pessoas dedicadas à avaliação, tempo

disponibilizado, conhecimentos e competências em avaliação, ...); a terceira incide em elementos da cultura organizacional particularmente influentes da avaliação (nomeadamente o suporte da liderança e a receptividade dos membros da organização à avaliação); e a terceira remete para fatores contingenciais internos comumente associados à capacidade avaliativa (dimensão financeira, antiguidade, papéis desempenhados).

A terceira, quarta e quinta dimensões remetem mais diretamente para as concepções e práticas de avaliação, em termos das suas motivações, propósitos, abordagens metodológicas e usos. A terceira - concepções, propósitos e motivações – centra-se nos fatores explanatórios e modos de adoção de práticas de avaliação (ímpeto internos/externo; adoção coerciva/normativa/mimética/voluntária; teorias explicativas) e na forma como eles são genericamente entendidos nas organizações. A quarta dimensão – atividades e processos de avaliação - incide na intensidade/abrangência das práticas de avaliação e no seu desenho metodológico (conteúdos/âmbitos de incidência, tipos de dados, agentes e participantes). Por fim, a quinta dimensão reenvia às consequências da avaliação e à forma como é usado o conhecimento produzida pela avaliação, contemplando os tipos de usos da avaliação (de processo/de resultado, instrumentais/não instrumentais) e, numa outra perspetiva, os tipos de prestação de contas (interna/descendente/ascendente).

Para traçar perfis de avaliação nas OTS comecei por recorrer a procedimentos estatísticos (partindo dos resultados do inquérito), no sentido de agrupar informação que me permitisse reduzir a complexidade dos dados. Partindo da seleção de conjunto de variáveis associadas às diversas dimensões de análise da cultura avaliativa (23, no total)¹⁷⁹,

¹⁷⁹ Variáveis selecionadas – **Contextos, enquadramentos e percepções sobre avaliação:** A avaliação é parte integrante das nossas práticas de gestão; A avaliação ajuda-nos a fazer escolhas estratégicas sobre o futuro da organização; A avaliação ajuda-nos a melhorar a qualidade dos serviços que oferecemos; Os nossos financiadores estão muito interessados na avaliação das intervenções; Ocupar tempo e recursos na avaliação tira-nos disponibilidade para o mais importante: prestar serviços; Não temos o conhecimento ou experiência necessários para fazer avaliação com qualidade; Os requisitos de avaliação são sobretudo obstáculos que os financiadores nos impõem; O tempo e o dinheiro que gastamos em avaliação não compensam; Muito do que fazemos para a avaliação dos projetos/atividades é apenas simbólico. **Propósitos/motivações:** [Fazemos avaliação...] Para atrair recursos e promoção perante financiadores; Para tomar decisões organizacionais; Porque os financiadores exigem; Para promoção perante a comunidade local. **Usos da avaliação:** Obter certificação/acreditação; Obter novos financiamentos; Para reportar aos órgãos sociais Para decisões sobre

recorri à Análise Categrial de Componentes Principais (ACCP), que me permitiu resumir a informação presente nas variáveis e explorar as suas correlações. Usando-se como critério de retenção a regra do *eigenvalue* superior a 1, foi possível resumir informação relacional entre variáveis em dez componentes ortogonais que explicam **68% da variância total** das variáveis originais. Na tabela 11 (vide anexo 5) reproduzem-se os “pesos” de cada variável em cada componente e a percentagem de variância explicada. De acordo com estes pesos foi possível nomear a primeira componente como a **“Avaliação enquanto ferramenta de gestão estratégica”** (que explica 15% da variância total), dados os pesos elevados das variáveis “A avaliação ajuda-nos a fazer escolhas estratégicas sobre o futuro da organização”, “A avaliação ajuda-nos a melhorar a qualidade dos serviços que oferecemos” e “Fazemos avaliação para nos ajudar a tomar decisões organizacionais”. A segunda componente, que explica 9% da variância total das variáveis originais, foi designada de **“Avaliação enquanto ferramenta de projeção externa”**, mediante o forte peso de variáveis como “Os nossos financiadores estão muito interessados na avaliação das intervenções” e “Fazemos avaliação para nos promovermos perante financiadores e atrair recursos”. A quarta, quinta e décima componentes (que globalmente explicam 16% da variância total) foram nomeadas de **“Avaliação como fardo ou dispersão”** dados os scores elevados das variáveis associadas a perceções negativas sobre a avaliação (por exemplo, “Os requisitos de avaliação são sobretudo obstáculos que os financiadores nos impõem” e/ou “Muito do que fazemos para a avaliação dos projetos/atividades é apenas simbólico”). Por fim, as componentes 3, 6, 7, 8 e 9 (que globalmente explicam 27% da variância total) foram designadas de **“Avaliação como mecanismo de responsabilização interna e promoção local”** já que, embora apresentem uma grande diversidade interna, possuem em comum o peso elevado das variáveis “Fazemos avaliação para promoção perante a comunidade local” e/ou “Usamos a avaliação para reportar aos órgãos sociais”.

Este exercício permitiu, assim, identificar lógicas reveladoras das grandes partições das visões e enquadramentos da avaliação nas organizações. Partindo destas grandes linhas, foram acrescentados elementos de caracterização a cada um dos perfis,

alocação de recursos; Para decisões sobre recursos humanos; Para reportar a financiadores; Para rever estratégias; Para divulgação e relações públicas; Para rever projetos; Não uso.

mediante a utilização dos *scores* associados às componentes extraídas na Análise Categorical de Componentes Principal para desenvolver múltiplas análises estatísticas. Com efeito, desenvolverem-se análises de correlação bivariada¹⁸⁰ entre cada um dos componentes da ACCP e um conjunto vasto de indicadores das **diversas dimensões e subdimensões de cultura avaliativa** exploradas no inquérito por questionário e cujos resultados podem ser consultados no anexo 5 (tabelas 12, 13 e 14). A composição dos perfis, sintetizada no quadro 18, foi alimentada pelos resultados das entrevistas, que permitiram uma análise mais detalhada, compreensiva e consubstanciada dos diversos elementos característicos de cada um dos perfis.

Quadro 18. Perfis de avaliação nas OTS

		AVALIAÇÃO ENQUANTO FERRAMENTA DE GESTÃO ESTRATÉGICA	AVALIAÇÃO ENQUANTO FERRAMENTA DE PROJEÇÃO EXTERNA	AVALIAÇÃO COMO FARDO OU DISPERSÃO	AVALIAÇÃO COMO MECANISMO DE RESPONSABILIZAÇÃO INTERNA E PROMOÇÃO LOCAL
CONCEÇÕES, PROPÓSITOS E MOTIVAÇÕES	Visão global	Visão da avaliação como ferramenta de apoio à decisão estratégia e melhoria da qualidade	Visão da avaliação como uma ferramenta de atração de recursos e promoção perante os financiadores	Visão da avaliação como distração da missão e desperdício ou fardo imposto pelos financiadores sem retorno	Visão da avaliação enquanto ferramenta de promoção na comunidade local e de prestação de contas “obrigatória”
	Impulso /modo de adoção	Interno: Adoção voluntária e dissociada do financiamento	Interno: Adoção mimética, associada ao financiamento, mas deliberada	Externo: Adoção coerciva/simbólica (teoria da agência) ou normativa (teoria institucional)	Adoção normativa (práticas informais) ou adoção coerciva (práticas formais)
	Propósito	Assegurar eficiência e qualidade <i>Output legitimacy</i>	Reduzir a incerteza, atrair recursos e assegurar qualidade	Garantir sobrevivência e obter legitimidade	Obter legitimidade

¹⁸⁰ No caso dos fatores contingenciais internos e de outras variáveis de caracterização organizacional (antiguidade, dimensão, tipo de organização, tipos de funções, subsectores de atividade, perfil do financiamento, formas e estatutos) optou por se conduzir análises multivariadas (regressão logística), dada a relevância de isolar os efeitos de cada uma delas dos efeitos das restantes.

		AVALIAÇÃO ENQUANTO FERRAMENTA DE GESTÃO ESTRATÉGICA	AVALIAÇÃO ENQUANTO FERRAMENTA DE PROJEÇÃO EXTERNA	AVALIAÇÃO COMO FARDO OU DISPERSÃO	MECANISMO RESPONSABILIZAÇÃO INTERNA E PROMOÇÃO LOCAL
ATIVIDADES E PROCESSOS	Práticas	Forte abrangência/intensidade das práticas de avaliação, sobretudo monitorização e avaliação formal à escala organizacional	Elevada amplitude e intensidade de atividades avaliativa, sobretudo associadas a projetos	Abrangência/intensidade das práticas de avaliação moderada; Forte associação à certificação, particularmente modo de adoção normativa	Dissociação da generalidade das práticas , à exceção dos relatórios anuais de atividades; Avaliação informal ou <i>ad hoc</i>
	Métodos	Diversidade de estratégias de recolha de dados; Prevalência da abordagem qualitativa	Múltiplas técnicas de recolha de informação, mas domínio do inquérito; Prevalência da abordagem quantitativa	Alguma diversidade metodológica, mas apenas associação forte com inquérito	Recolha casual de informação: conversas informais/impressões; Dissociação dos métodos de recolha de informação primária.
	Conteúdos	Satisfação e Efeitos	Efeitos	Satisfação	Realizações
	Agentes	Avaliação mista	Avaliação externa ou mista	Avaliação externa ou mista	Autoavaliação informal
	Participação	Elevada participação interna; Participação externa circunscrita a utilizadores/as	Elevada participação Forte envolvimento das entidades financiadoras; Alheamento dos órgãos de governação	Baixos níveis de participação Maior participação interna, no modelo normativo e externa, no modelo coercivo; Alheamento dos órgãos de governação	Fraca participação externa; Participação restringe-se a órgãos de governação
USOS	Tipos de Uso	Uso instrumental (e exclusivamente instrumental); Dissociação do uso não instrumental Foco no uso da avaliação para tomada de decisão	Uso misto, mas sobretudo não instrumental; Utilidade instrumental centrada na gestão de projetos	Não uso ou uso não instrumental; Dissociação dos usos mistos	Padrões diversos, mas prevalência dos usos não instrumentais
	Prestação de contas	Prestação de contas interna e descendente; Destinatários principais: órgãos de governação e utilizadores/as	Prestação de contas alargada (interna, ascendente e descendente); Destinatários principais: financiadores e parceiros	Prestação de contas restrita: consumo da equipa técnica e operacional	Prestação de contas restringe-se a órgãos sociais e associados

		AVALIAÇÃO ENQUANTO FERRAMENTA DE GESTÃO ESTRATÉGICA	AVALIAÇÃO ENQUANTO FERRAMENTA DE PROJEÇÃO EXTERNA	AVALIAÇÃO COMO FARDO OU DISPERSÃO	AVALIAÇÃO COMO MECANISMO RESPONSABILIZAÇÃO INTERNA E PROMOÇÃO LOCAL
CONTEXTO ORGANIZACIONAL INTERNO	Estruturas	Enquadramento da avaliação em iniciativas de gestão mais amplas (PE, GQ); Estruturas formais alinham-se com as necessidades da organização	Enquadramento em projetos e processos de certificação e/ou PE; Institucionalizaçã o formal e informal	Enquadramento em processos de certificação e/ou projetos; Institucionalização formal: avaliação integrante das práticas de gestão	Inexistentes.
	Recursos	Recursos financeiros, técnicos e de tempo alocados à avaliação considerados suficientes; Diversidade de fonte de recursos para avaliação	Avaliação suportada por fundos externos (sobretudo projetos); Fortes competências internas em avaliação	Recursos financeiros, de tempo e de competências não são problema	Avaliação informal, sem custos ou sustentada por fundos internos; Sem competências
	Cultura organizacional	Forte cultura de aprendizagem: suporte e comprometimento da liderança da organização; recetividade à avaliação; forte visão da avaliação como oportunidade de aprendizagem	Cultura de aprendizagem moderada pelas exigências de prestação de contas; Forte recetividade da avaliação por parte dos membros da organização	Fraca cultura de aprendizagem: perceção da avaliação como uma ameaça por parte dos membros da organização	Fraca cultura de aprendizagem: perceção da avaliação como uma ameaça por parte dos membros da organização; Falta de suporte da liderança
	Fatores contingenciais	Funções mistas; Grande dimensão financeira	Funções mistas ou apenas de serviço; Menor antiguidade	Funções só de expressão ou só de serviço; Maior antiguidade	Menor dimensão; Maior antiguidade
CONTEXTO ORGANIZ. EXTERNO	Pressões/expetativas do campo organizacional	Fraca pressão externa; Dissociação do financiamento por projeto	Forte pressão externa; Forte pertença a organismos/rede s internacionais; Financiadores interessados na avaliação; Apoio financeiro e técnico à avaliação	Forte pressão externa, mas sem retorno dos financiadores; Falta de apoio por parte das entidades financiadoras/ tutelares	Fraca pressão externa; Dissociação dos projetos; Reduzido interesse dos financiadores pela avaliação

	AVALIAÇÃO ENQUANTO FERRAMENTA DE GESTÃO ESTRATÉGICA	AVALIAÇÃO ENQUANTO FERRAMENTA DE PROJEÇÃO EXTERNA	AVALIAÇÃO COMO FARDO OU DISPERSÃO	AVALIAÇÃO COMO MECANISMO RESPONSABILIZAÇÃO INTERNA E PROMOÇÃO LOCAL
Perfil do financiamento	Entidades não dependentes do financiamento público (forte peso da filantropia e de receitas próprias)	Financiamento por projeto	Fraco peso da filantropia na estrutura de receitas	Forte peso das receitas próprias
Formas, estatutos e subsectores de atividade	Forte associação à detenção de estatuto de IPSS, mas não dependentes do financiamento público	Forte dissociação do estatuto de IPSS: prevalência de ADL e ONG de setores diversos	Prevalência dos setores da participação cívica e do associativismo profissional	Prevalência das coletividades, mas também um peso expressivo de misericórdias e cooperativas de ensino

Fonte: Elaboração da autora.

Perfil I – Avaliação enquanto ferramenta de gestão estratégica

Este perfil de avaliação integra uma visão globalmente **positiva** sobre a avaliação, quer em termos de implementação quer em termos dos seus benefícios para a organização. A avaliação é aqui entendida, essencialmente, como uma ferramenta de apoio à decisão estratégica e à melhoria da qualidade da gestão e da intervenção da organização, com um claro interesse no **valor racional e objetivo da avaliação**.

A noção de “organização de ação”, descrita por Brunsson (2002) e retomada por Højlund (2014) para caracterizar o modo de adoção voluntário das práticas de avaliação, ajusta-se ao perfil. O impulso para a avaliação surge particularmente vinculado a **fatores internos** e a uma “**responsabilidade sentida**” (Fry, 2005), ancorada em valores e numa dimensão ética. As práticas de avaliação são adotadas **voluntariamente**, por organizações com elevada propensão interna para avaliar, para colmatar as necessidades de conhecimento, com o fim essencial de melhorar o desempenho. Aliás, a característica que melhor distingue este perfil do seguinte é a **dissociação da avaliação da pressão externa** exercida por forças ligadas ao financiamento, o que é propiciado pelas características da sua estrutura de receitas – robusta, estável, diversificada em termos de fontes e independente de subsídios públicos. A avaliação é conduzida essencialmente com o intuito de informar a tomada de decisão, em resposta às necessidades internas de gestão (tal

como previsto pela teoria da escolha racional), mas também para adquirir "output legitimacy" (legitimidade pelos resultados) (Schmidt, 2013) pelo bom desempenho das atribuições. Em todo o caso, esta legitimidade ancora-se na garantia da qualidade da intervenção e na prossecução da missão e não na procura de sobrevivência.

A utilidade da avaliação é, pois, sobretudo **instrumental**: para ser mobilizada para melhorar a qualidade dos serviços, para fazer escolhas estratégicas sobre o futuro da organização ou para tomar outras decisões organizacionais (alocação de recursos humanos, financeiros e materiais). É, aliás, o único perfil em que se verifica uma associação positiva estatisticamente significativa com os usos exclusivamente instrumentais. Existe, pois, um forte sentido de racionalidade e operacionalidade nos **benefícios de aprendizagem** associados a este perfil, em que a avaliação é usada como base para a ação e para a mudança mediante a implementação das respetivas recomendações. As dinâmicas de **prestação de contas** vinculadas a estes efeitos são sobretudo **internas e descendentes**: responsabilização interna perante a missão e pessoal – que inclui decisores e pessoal do terreno –, e responsabilização descendente perante utilizadores/as e comunidade local. A relação positiva forte deste perfil com **o consumo regular da informação avaliativa por parte dos órgãos de direção estatutários** constitui, aliás, uma tendência que o distingue dos restantes perfis com práticas de avaliação significativas (isto é, em que a associação às práticas de avaliação é positiva).

Este é o perfil em que os exercícios de avaliação são mais intensamente **participados internamente** – pela equipa técnica/operacional e pela direção estatutária e técnica. A associação com a participação externa é menos marcada, ainda que significativamente positiva, especialmente a que envolve utilizadores/as ou beneficiários/as.

Por seu turno, a intensidade e o tipo de participação refletem, em parte, o tipo de práticas de gestão e avaliação privilegiadas. Este perfil é, desde logo, aquele em que a abrangência/intensidade das **práticas específicas de avaliação e medição do desempenho é superior**¹⁸¹, sobretudo as relativas a exercícios de monitorização e avaliação formal à

¹⁸¹ Para medir a amplitude das práticas de gestão e avaliação foram criados **dois scores**: um da **generalidade das práticas de gestão e avaliações** listadas no inquérito – frequentemente identificadas nas organizações entrevistadas com práticas de monitorização e avaliação –, que incluem os sistemas de gestão

escala organizacional. A associação às práticas de gestão e controlo é também forte, com particular incidência na avaliação de desempenho individual (que foi integrada no *score* global das práticas de gestão e não no específico das práticas de avaliação). Dentre a diversidade de exercícios de gestão e controlo elencados, os relativos à monitorização/auditora externa e ao reporte às entidades financiadoras, vinculados à prestação de contas ascendente, são os que se associam de forma menos significativa a este perfil, o que justificará a afirmação do fraco envolvimento das entidades financiadoras/tutelares nos exercícios avaliativos.

Um outro padrão de comportamento associado a este perfil (embora não exclusivo, como veremos), é comum ao identificado por Carman (2010) no seu estudo relativamente ao grupo de OTS com maior capacidade avaliativa, e respeita a tendência para o **enquadramento da avaliação em iniciativas de gestão mais amplas**, nomeadamente em sistemas de gestão da qualidade e/ou de planeamento estratégico. No primeiro caso, a informação de avaliação e medição do desempenho é usada para monitorizar o progresso (trimestral/quadrimestral/semestral e/ou anual) dos objetivos estratégicos definidos. No segundo caso, a avaliação é perspetivada como parte de uma iniciativa mais abrangente de garantia/melhoria da qualidade, na qual a pessoa responsável pelo sistema de gestão da qualidade é também quem gere as atividades de avaliação.

Estes processos mais amplos conferem à avaliação uma **estrutura mais formalizada** que, no caso deste perfil, se **alinha com as necessidades e propósitos de melhoria da organização**. Ou seja, as estruturas formais de avaliação – que aqui se

da qualidade, o planeamento estratégico, as auditorias, a avaliação do pessoal e o reporte anual de atividades e sobre execução física e financeira de projetos, e **outro** incluindo apenas **práticas específicas de monitorização/medição do desempenho e avaliação** (concretamente: monitorização do cumprimento de objetivos e metas organizacionais; avaliações formais do desempenho organizações; avaliações formais de projetos; modelos lógicos; análises custo-benefício do valor social gerado e *balanced scorecard*). Os *scores* foram calculados para cada uma das organizações inquiridas a partir dos níveis de envolvimento (0 - Nunca, 1 – Ocasionalmente, 3 – Regulamente) em cada uma das atividades selecionados no inquérito. O nível/score de práticas global (incluindo todas as práticas listadas) podem variar desde 0 (nenhum envolvimento em nenhuma das práticas de gestão e avaliação listadas) até 28 (envolvimento regular em todas as práticas), enquanto o nível/score de práticas de monitorização e avaliação pode variar entre 0 (nenhum envolvimento em nenhuma das práticas de monitorização e avaliação listadas) até 12 (envolvimento regular em todas as práticas).

apresentam sobretudo à escala organizacional (e não de projetos) – refletem não apenas um processo de institucionalização formal, mas também informal, assente numa **forte cultura de aprendizagem**. A avaliação é, aqui, parte de um processo mais amplo de aprendizagem, sendo institucionalizado como parte do sistema de informação organizacional, estrutura de poder, processos e sistemas que influenciam a tomada de decisão e a ação (Boyle, Lemaire e Rist, 1999). O suporte e comprometimento da liderança da organização, a receptividade da avaliação por parte do pessoal da organização e a forte visão da avaliação como oportunidade de aprendizagem, compõem o retrato de uma cultura organizacional que suporta a avaliação, confirmando as asserções da alínea f) da minha primeira hipótese.

O **grande volume de atividades de recolha de informação** e a **amplitude dos tipos de informação recolhida** emerge como um padrão distintivo deste perfil. Tal diversidade é coerente com a multiplicidade de práticas de avaliação (e de gestão, em geral) e a intensidade do envolvimento dos *stakeholders* (sobretudo internos) nos exercícios avaliativos caracterizadora deste perfil. As estratégias de recolha de dados compreendem técnicas qualitativas e quantitativas, embora os métodos conotados com as abordagens qualitativas pareçam prevalecer, mediante associações mais fortes com técnicas como a observação direta, os grupos de trabalho sobre avaliação e as entrevistas focalizadas individuais e de grupo. O foco dos exercícios avaliativos, em termos de conteúdos, é também diversificado, incidindo sobre os vários níveis da cadeia de resultados das intervenções – desde as realizações (*outputs* físicos e financeiros) até aos impactos. Não obstante, a prioridade dos exercícios de avaliação é recolher informação sobre a satisfação de quem beneficia dos serviços ou participa nas atividades e sobre os efeitos produzidos na vida das pessoas, em detrimento da informação sobre *outputs* gerados pelas atividades (n.º de ações, n.º de participantes, montante gasto, etc.).

Neste perfil, a solidez da cultura de avaliação encontra-se alicerçada na disponibilidade de recursos, quer financeiros e materiais, quer de conhecimentos e competências necessários para conduzir ou administrar atividades de avaliação. A avaliação tende a ser conduzida por pessoas com competências específicas em avaliação, combinando o recurso à consultoria externa e ao pessoal interno não diretamente

responsável pela intervenção a ser avaliada (em geral, a equipa de gestão/gestão da qualidade) configurando uma modalidade de **avaliação mista**. Apesar do recurso significativo à consultoria externa, os custos com a avaliação são suportados no quadro das **despesas administrativas gerais**.

As organizações de grande dimensão financeira, com estatuto de IPSS mas com uma estrutura de receitas diversificada e não dependente de subsídios públicos, e com um perfil de funções híbrido, são as que melhor encaixam neste perfil.

Perfil II – Avaliação enquanto ferramenta de projeção externa

A avaliação é percebida, neste perfil, como uma ferramenta de atração de recursos e promoção perante os financiadores. Enquanto no perfil anterior o foco era colocado no valor racional e instrumental da avaliação, aqui é colocada a tónica no seu **valor concetual e simbólico**.

Enquanto no anterior perfil a avaliação se dissociava das forças externas, neste parece existir uma **pressão forte do campo organizacional** para a condução de exercícios avaliativos. Mas, como este tipo de organização é recetivo às ideias e conceitos de avaliação, as práticas são integradas de forma deliberada e não por coerção, procurando estrategicamente “reproduzir” as melhores práticas do campo organizacional para aprender e se tornar mais eficiente. As possibilidades de identificação de organizações “bem-sucedidas” são potenciadas pela pertença a **redes** de organizações (nomeadamente, à **escala internacional**), a que este perfil se associa de forma estatisticamente significativa.

As entidades que se aproximam deste perfil tendem a operar num **ambiente organizacional incerto** onde, por exemplo, não existe uma base estável de recursos, nomeadamente quando a sobrevivência das organizações se encontra financeiramente dependente da aprovação de projetos, funcionando “por projeto”. Existe, efetivamente, uma associação estatisticamente significativa deste perfil com o **financiamento assente em projetos** e, concomitantemente, uma dissociação significativa do estatuto de IPSS, o qual, condicionando o estabelecimento de acordos de cooperação com a Segurança Social, contribui para a estabilidade de receitas das organizações (Martín *et al.*, 2014).

Ao contrário do perfil anterior, em que a avaliação é sobretudo usada para informar a tomada de decisão, neste perfil, ela é usada essencialmente para atrair recursos

e alavancar a obtenção de novos financiamentos. Os mecanismos de responsabilização (relatórios de execução física e financeira, avaliações formais e/ou auditorias) são, pois, usados estrategicamente como meios para receber algo em troca (e não apenas para prestar contas), tal como preconizado **pela teoria da dependência de recursos**. Um elemento distintivo deste perfil relativamente ao anterior respeita, exatamente, à elevada relevância e profusão de **usos não instrumentais** e à maior convivência de usos instrumentais e não instrumentais (**usos mistos**). Com efeito, a par do uso instrumental para **planear/rever estratégias** (sobretudo associadas a projetos), a avaliação é usada conceitualmente e/ou simbolicamente para responder a requisitos de financiamento, legitimar a intervenção e/ou **assegurar ou atrair novos financiamentos** – “responder às matrizes de avaliação das candidaturas”, “justificar projetos seguintes”, “fortalecer as novas candidaturas”, “demonstrar às entidades financiadoras que o dinheiro que aplicaram foi corretamente aplicado”, “visibilizar o trabalho”, “sensibilizar para o trabalho realizado” “demonstrar que a instituição já avalia o seu impacto”, são referências colhidas em entrevistas. As características deste perfil, no que toca às relações de *accountability*, alinham-se com os pressupostos da alínea b) da segunda hipótese do estudo.

Os **benefícios da aprendizagem** associados a este perfil colocam-se a um nível mais abrangente e tendencialmente menos operacional que no anterior, situando-se primordialmente ao nível da mudança nos entendimentos e atitudes das partes interessadas, suscitados pela sensibilização sobre aspetos operacionais ou sobre os resultados das intervenções. A **prestação de contas** associada a estes benefícios é **alargada** (interna, ascendente e descendente), aproximando-se da **abordagem dos stakeholders** à responsabilização, no seu olhar integrado sobre a forma com as organizações lidam com relações e exigências múltiplas. Com efeito, além da responsabilização interna perante os membros da organização, e descendente perante parceiros, comunidade local e utilizadores/as, este perfil associa-se fortemente à responsabilização ascendente perante entidades financiadoras/tutelares. Por outro lado, ao contrário do perfil anterior, neste os órgãos de decisão estatutários tendem a dissociar-se do consumo regular da informação avaliativa, o qual tende a concentrar-se na direção técnica/executiva e outro pessoal técnico.

Este é o perfil em que os exercícios de avaliação são, globalmente, **mais intensamente participados**. A intensidade da participação é particularmente devedora do envolvimento de **atores externos** às organizações, sobretudo representantes das entidades financiadoras e públicos-alvo, mas também a comunidade local. Além da centralidade da participação externa neste modelo, um elemento distintivo relativamente ao modelo anterior respeita ao (não) envolvimento dos órgãos de governação. Observamos, pois, neste perfil, que quer a direção quer os restantes órgãos sociais tendem a dissociar-se da participação nos exercícios avaliativos, em favor da direção técnica/executiva e restante equipa, o que aponta para uma forte profissionalização da avaliação neste modelo.

O padrão de envolvimento dos *stakeholders* associado a este perfil reflete também o tipo de práticas de avaliação (e de gestão, mais genericamente) adotadas. Tal como no perfil anterior, verifica-se uma tendência forte para a adesão a práticas de gestão e avaliação. Não obstante, existe uma maior associação positiva ao *score* de práticas de gestão do que ao *score* das práticas específicas de monitorização e avaliação (ao contrário do que acontecia no perfil anterior, em que se verificavam níveis de associação aproximados). Tal deve-se à maior amplitude de práticas de controlo e gestão a que este modelo se associa e das quais o modelo anterior se distanciava, ou se ligava de forma mais moderada. É o caso das práticas de **relatório descritivo** (execução física e financeira) e de **monitoria/auditoria externa** conotados com o desenvolvimento de **projetos**, às quais este perfil se associa de forma paradigmática, mas que não identificam o anterior modelo. Depois, são exatamente as práticas de **avaliação formal** associadas aos **projetos** aquelas que catapultam o *score* de avaliação neste perfil, no qual as práticas de avaliação formal à escala organizacional são menos proeminentes. A forte identificação com este tipo de práticas, particularmente conotadas com a prestação de contas ascendente, alinha-se com a afirmação do forte envolvimento das entidades financiadoras/tutelares nos exercícios avaliativos.

Apesar dos diferentes padrões observados ao nível das práticas de avaliação, em particular no que respeita ao nível de análise privilegiado (organização como um todo e projetos, respetivamente), à semelhança do anterior, também neste perfil se verifica a

tendência para a **articulação da avaliação com iniciativas de gestão estratégica e da qualidade**: planeamento estratégico e certificação da qualidade. Esta articulação permite uma maior integração da informação de avaliação associada aos projetos nos processos de planeamento e tomada de decisão de nível organizacional. Concorre também para conferir uma estrutura mais formalizada às atividades de avaliação conduzidas fora do âmbito dos projetos, embora neste perfil a institucionalização da avaliação esteja particularmente associada à dinâmica dos projetos e seja **contingente aos requisitos de financiamento** (nomeadamente em termos de regras, orientações e calendarização).

As contingências associadas ao financiamento por projeto e respetivos requisitos de prestação de contas ascendentes atuam, este perfil, como **mediadoras da cultura de aprendizagem**. Mas neste modelo, ao contrário do seguinte, parece estabelecer-se um **equilíbrio entre os papéis de accountability e de aprendizagem organizacional**. É que, mesmo sendo a avaliação um requisito necessário, as organizações com este perfil entendem que o **investimento em avaliação pode trazer algum retorno, quer em termos de eficiência quer em termos de atração de recursos**. Esta visão é sustentada por uma cultura organizacional que promove a reflexão e que se caracteriza pela abertura dos seus membros à aprendizagem.

À semelhança do perfil anterior, também este se caracteriza pela **amplitude de conteúdos/âmbitos de incidência** da avaliação em termos de tipos de informação recolhida, incidindo o sobre os vários níveis da cadeia de resultados das intervenções, ainda que a prioridade seja a recolha de dados sobre os **impactos/mudanças** suscitados pelas intervenções. Já no que toca aos métodos de avaliação, encontramos neste perfil um maior afunilamento das abordagens adotadas em torno da técnica de **inquérito** por questionário, fazendo prevalecer a perspetiva quantitativa.

Neste modelo, a solidez da cultura de avaliação encontra-se também sustentada na **disponibilidade de recursos financeiros, humanos e técnicos** necessários à dinamização de exercícios de avaliação. Mas, ao contrário do primeiro modelo, neste a avaliação tende a ser financiada por **fundos externos**, essencialmente no âmbito de projetos. Também o recursos à **avaliação externa** – que aqui tanto pode ser entendida como da iniciativa das

organizações (que contratam serviços externo para a condução das avaliações) como das entidades financiadoras/tutelares (no caso das auditorias) – tende a ser mais frequente neste perfil.

As organizações de constituição mais recente, ONG ou ADL de diversos setores atividade e perfil de funções misto ou só de serviço são as que mais se associam a neste perfil.

Perfil III – Avaliação como fardo ou dispersão

Este perfil, embora partilhe características com os dois anteriores, diverge essencialmente na forma como a avaliação é percebida e enquadrada. A avaliação é tendencialmente concebida neste modelo – mais **heterogéneo** que os anteriores¹⁸² – como um **fardo imposto pelas entidades financiadoras/tutelares** ou como um **desvio da missão** das organizações.

O padrão que mais identifica este modelo e o distingue dos anteriores respeita à **utilidade** dos exercícios de monitorização e avaliação. Com efeito, este perfil caracteriza-se, grosso modo, pela **perceção de inutilidade** dos exercícios avaliativos e pela associação forte ao **não uso** da avaliação. Apesar da **presença de estruturas formais de avaliação** e de ela estar **incorporada no trabalho regular destas organizações** – o que é atestado pela abrangência /intensidade razoável das práticas de avaliação (e das práticas de controlo e gestão em geral)¹⁸³ e pela forte adesão à condição da avaliação enquanto “parte integrante das práticas de gestão” –, paradoxalmente, elas não procuram utilizar as avaliações de forma instrumental.

A dissociação dos usos da avaliação para aprendizagem organizacional encontra-se ancorada no impulso para adoção deste tipo de práticas. Neste perfil, o **impulso** para a avaliação é eminentemente **externo**: resultado de **pressões coercivas** ou **normativas** no ambiente. A avaliação tende a ser conduzida nas organizações por razões **simbólicas ou cerimoniais**, com o objetivo de serem percebidas como legítimas e bem-sucedidas

¹⁸² Enquanto os anteriores perfis foram compostos a partir de apenas uma componente da ACCP, este perfil (assim como o que se segue) é definido por mais do que uma componente (que adiante designo também de “segmento”).

¹⁸³ A força das associações ao *score* de gestão e ao *score* de avaliação é semelhante.

(conforme pressuposto **da teoria institucional**)¹⁸⁴, ou simplesmente porque a avaliação lhes é **imposta** pelas entidades financiadoras (na linha dos preceitos da **teoria da agência**)¹⁸⁵. Este tipo de organização **adotará práticas de avaliação**, mas tenderá as **desconectá-las (*decouple*)**¹⁸⁶ **dos processos normais da organização**, nomeadamente da tomada de decisão. Segundo Meyer e Rowan (1977), as pressões institucionais podem contradizer as necessidades internas de eficiência. As organizações podem, então, tentar adaptar-se, mas dissociando a sua ação rotineira da estrutura/estratégia formal, no sentido de preservar a eficiência organizacional e, ao mesmo tempo, **obter legitimidade**. Esta estratégia pode, então, ser vista como uma forma de gerir o dilema de resistir à monitorização externa, conjugado com propósito de ganhar legitimidade a partir do mesmo contexto externo (Arvidson e Lyon, 2014: 871). Neste perfil, a procura de legitimidade e reconhecimento externo é feita, tendencialmente, mediante a **obtenção de selo de certificação** de serviços/produtos. A certificação, a que este perfil se associa fortemente, é aqui usada de forma essencialmente **simbólica**, para demonstrar legitimidade e reforçar a reputação perante financiadores e comunidade, **e não para a melhoria** dos processos internos e da intervenção juntos dos públicos-alvo, como acontecia nos anteriores modelos.

O risco é que, como bem nota Carman (2011: 365), as pressões normativas e coercitivas empurrem as organizações para a adoção de comportamentos ritualísticos que, em última análise, desvalorizam as avaliações e dão credibilidade à crença de que a avaliação subtrai recursos às organizações que poderiam ser melhor utilizados para melhorar as intervenções.

É também no enquadramento deste perfil que encontramos o maior potencial para a **manipulação ou distorção** da informação avaliativa (“dar um jeitinho para chegar à meta”, conforme nos reporta um técnico de uma OTS), particularmente nas situações em

¹⁸⁴ Neste caso, os exercícios de avaliação são perspetivados enquanto “padrões do setor” ou normas profissionais (DiMaggio e Powell, 1983) – algo em que é suposto as organizações se envolverem se querem ser percebidas como legítimas (Carman, 2011: 355).

¹⁸⁵ Neste caso, a avaliação é conduzida para efeitos de monitorização ou cumprimento de requisitos externos.

¹⁸⁶ O conceito de dissociação (*decoupling*) descreve a forma como as organizações “tornam visíveis os compromissos públicos para satisfazer as exigências de seu ambiente externo, enquanto as operações internas da organização podem seguir trajetórias diferentes” (Arvidson e Lyon, 2014: 871).

que as avaliações têm um valor **punitivo** (nomeadamente mediante cortes no financiamento) e/ou são determinadas e formatadas tendo sobretudo ou apenas em conta necessidades dos financiadores, consideradas sem sentido ou inapropriadas e prejudiciais. Aqui é colocado a nu o conflito de prioridades, objetivos e valores entre organizações (enquanto agentes) e entidades financiadoras (enquanto principais), na linha do preconizado pela **teoria da agência**.

Não obstante, se a perspetiva de sanções na sequência de exercícios de avaliação pode fomentar fenómenos de empolamento ou ocultação de realizações e resultados e a associar-se a visões negativas da avaliação, a **perceção de que o investimento em avaliação** (em termos dos seus processos e resultados) **não é reconhecido pelas entidades financiadoras/tutelares** é claramente **dissuasora** deste tipo de investimento e precursora da perceção da avaliação como **inútil/sem retorno**, elevando o seu carácter simbólico e cerimonial. De resto, a importância do retorno externo do investimento na avaliação surge como central não só no discurso de representantes de organizações de base, mas também de organismos representativos do setor e de especialistas/profissionais de avaliação.

O desinteresse ou não uso da avaliação por parte das entidades financiadoras reflete uma **dualidade nas relações de accountability** já evidenciada em análises anteriores (por exemplo, Fine, Thayer e Coghlan, 1998; Tassie, Murray e Cutt, 1998; Bickel, Millett e Nelson, 2003; Carman, 2009). Por um lado, os requisitos de financiamento prescrevem o envolvimento das OTS numa série de práticas para evidenciar empiricamente o desempenho. Por outro lado, as entidades financiadoras não são eficazes no uso dos dados de avaliação nos processos de tomada de decisão. Os requisitos de avaliação são, neste enquadramento, entendidos pelas organizações como inconsequentes e, portanto, não justificam investimento significativo de tempo e recursos, além do estritamente necessário para o cumprimento do que é exigido. As características deste perfil, em termos de relações de *accountability*, alinham-se com os pressupostos da alínea a) da segunda hipótese do estudo.

Em face da falta de suporte técnico e financeiro e de retorno do investimento em avaliação por parte das entidades financiadoras, parece observar-se, neste perfil, uma maior orientação para o reforço da legitimidade junto de outros *stakeholders*,

nomeadamente os **públicos-alvo** e a **comunidade local**. Esta orientação é visível na maior tendência para vincular a adoção de práticas de avaliação ao propósito de promoção perante a comunidade local, face ao propósito de promoção perante as entidades financiadoras (prevalente no perfil II) e ao propósito instrumental de apoio à tomada de decisões (prevalente no perfil I).

Às características de um ambiente externo, pouco propício à capitalização/maximização do potencial da avaliação, juntam-se algumas especificidades do **contexto organizacional interno**. Os maiores constrangimentos a este nível não se colocam tanto em termos de recursos - financeiros, materiais ou técnicos -, mas mais em termos de cultura organizacional. Com efeito, este modelo parece convergir para uma fraca cultura de aprendizagem assente, nomeadamente, na perceção da avaliação como uma ameaça por parte dos membros da organização e na ausência de uma atitude partilhada que considere a avaliação uma oportunidade de aprendizagem.

As práticas de avaliação conduzidas neste contexto caracterizam-se por **moderados ou baixos níveis de envolvimento dos stakeholders**. A abrangência da participação na conceção, implementação e consumo dos exercícios de avaliação é genericamente reduzida, embora se verifiquem associações positivas estatisticamente significativas (embora fracas) com o envolvimento de certos grupos de *stakeholders*. O estabelecimento de um padrão comum de envolvimento é contudo desafiado pela diversidade interna do perfil. Não obstante, é possível assinalar uma tendência comum: a fraca associação ao envolvimento dos órgãos de governação estatutários nestes exercícios, quer na implementação quer no consumo da informação avaliativa.

Em termos de estratégias de recolha de dados de avaliação este perfil é também menos rico e abrangente que os anteriores: o **peso da informação secundária** no sistema de informação é **superior** e a recolha de dados primários é feita essencialmente através de **inquérito por questionário**. A associação aos diversos tipos de dados também é menor neste perfil que, embora mais heterogéneo internamente, em termos globais tende a priorizar a informação sobre **satisfação** de beneficiários/as e/ou outros *stakeholders*. Outro aspeto transversal a este modelo é a associação à recolha de informação sobre padrões de referência/ *benchmarks* (referência ou medida padrão para comparação), o que ilustra a

preocupação, no seio deste modelo, em fazer comparações com outras organizações, ou a identificação de medidas/métricas de medição do desempenho/impacto comuns.

O tipo de organização exclusivamente prestadora de serviços ou com funções puramente expressivas, dos setores da participação cívica e do associativismo empresarial, é prevaletentes neste perfil.

Perfil IV – Avaliação como mecanismo de responsabilização interna e promoção local

A característica essencial que distingue este perfil dos anteriores tem que ver com o nível de envolvimento em práticas de avaliação: enquanto os anteriores modelos se caracterizam pelo envolvimento forte ou moderado em atividades de avaliação, este destaca-se pelo **envolvimento fraco ou nulo em práticas de avaliação**.

Vemos que este modelo, embora bastante heterogéneo, globalmente tende a dissociar-se da generalidade das práticas de gestão e avaliação listadas no inquérito, à exceção daquelas a que as organizações estão estatutariamente obrigadas: os relatórios anuais de contas e atividade¹⁸⁷. A (auto)avaliação tende a assumir um **carácter informal** (causal, espontâneo), **esporádico ou ad hoc**. A informalidade e carácter pontual dos exercícios avaliativos conduzidos neste modelo estão refletidos na **ausência de estruturas formais** e de **metodologias de observação sistemáticas**. A recolha de informação é casual e tem por base impressões ou o *feedback* esporádico dos públicos-alvo. Em face da ausência de estratégias de recolha de informação primária, os dados que sustentam a avaliação são tendencialmente **numéricos**, incidido sobretudo em aspetos financeiros (receitas e despesas).

A debilidade metodológica neste perfil é também revelada pelo fraco envolvimento das partes interessadas nos exercícios de avaliação, em especial as externas às organizações. A participação na condução dos exercícios de avaliação tende a restringir-se aos **órgãos sociais**, e a divulgação dos respetivos resultados confina-se a estes órgãos de gestão e outros membros associados e, nalguns segmentos (essencialmente associações de cultura, desporto e recreação), à **comunidade local**. Os principais propósitos que

¹⁸⁷ Apesar de esta prática apresentar associações globalmente mais fortes com os segmentos/componentes que compõem este perfil quando comparada com as correlações estabelecidas com as restantes atividades, a força das associações é fraca.

animam a adoção das práticas de avaliação neste modelo são, de resto, coerentes com estes padrões de envolvimento: a **promoção na comunidade local** e/ou a de **prestação de contas aos “stakeholders obrigatórios”** (essencialmente membros associados ou cooperantes e, em alguns segmentos, entidades financiadoras).

Neste perfil, os constrangimentos à avaliação são múltiplos, mas não são exclusivos da atividade de avaliação. Tal como Carman (2010) observa no seu estudo, relativamente ao grupo de organizações que enfrentam os maiores problemas na implementação de exercícios de avaliação, a avaliação é **só mais um dos muitos constrangimentos de gestão**. Também neste estudo, em alguns segmentos associados a este perfil, o escasso investimento na avaliação sinaliza um **problema de capacidade mais amplo**, essencialmente vinculado à **escassez de recursos** materiais, humanos e técnicos decorrentes da debilidade/instabilidade financeira destas organizações. A este segmento associam-se sobretudo entidades de menor dimensão financeira, essencialmente coletividades de cultura, desporto e recreio, mas também associações/cooperativas do setor da educação. Dentre as limitações associadas aos recursos, a **escassez de conhecimentos e competências** em avaliação é a mais proeminente, estendendo-se às entidades com maior dimensão financeira que se associam ao perfil – é o caso de algumas misericórdias e cooperativas de ensino.

O contexto interno neste modelo tende também a ser marcado por uma **cultura organizacional hostil à avaliação**, assente quer numa liderança pouco aberta e/ou comprometida com processos de reflexão, análise e aprendizagem contínua – especialmente em entidades que neste perfil se associam a uma constituição mais antiga – , quer numa atitude geral reativa à avaliação por parte dos membros da avaliação (perceção da avaliação como uma ameaça ou como uma perda de tempo).

Apesar da fraca adesão às práticas de avaliação e à multiplicidade de constrangimentos à sua implementação, este perfil caracteriza-se pelo convívio de visões positivas e negativas sobre avaliação. As visões positivas, tendencialmente mais comuns entre as coletividades, encontram-se vinculadas ao potencial a avaliação para **projeção da imagem e aproximação à comunidade local**, e as negativas, mais comuns nos setores da

educação, saúde e ação social, às **forças coercivas** do financiamento e à ausência de uma **cultura de aprendizagem**.

As organizações de grande dimensão financeira, com estatuto de IPSS mas com uma estrutura de receitas diversificada e não dependente de subsídios públicos e com um perfil de funções híbrido são as que melhor encaixam neste perfil.

Capítulo 6.2. Práticas heterogêneas em ambientes diversos: caminhos para o desenvolvimento de uma cultura avaliativa

Refreando os pressupostos de racionalidade que têm regido muita da discussão teórica e metodológica sobre avaliação, e que a descrevem como **instrumento racional e técnico usado internamente para tomar decisões**, o estudo revela que as motivações para a adoção de práticas de avaliação nas organizações e os usos que se fazem do seus resultados são muito diversos, e não se confinam ou esgotam naquela funcionalidade. Embora o pano de fundo do discurso das organizações a propósito da avaliação seja o da **internalização do “mantra” da avaliação** – a adesão à ideia da desejabilidade da avaliação e da sua utilidade instrumental – uma análise mais fina das práticas das organizações e da sua relação com variáveis do ambiente institucional deixam transparecer a prevalência das forças externas na adesão à avaliação, tal como coloquei em hipótese (alínea b) da primeira hipótese do estudo). Observa-se, coerentemente com pressupostos da teoria da agência e da dependência de recursos, que as **forças externas associadas ao financiamento exercem algum nível de controlo sobre a adesão à avaliação e tipo de práticas de avaliação adotadas**. O estudo mostra que o perfil do financiamento das organizações influencia de forma estatisticamente significativa quer a intensidade/abrangência das práticas de monitorização e avaliação quer as modalidades de avaliação adotadas em termos de metodologias e conteúdos. Independentemente da dimensão, antiguidade, perfil de atividade ou forma jurídica, as entidades cuja principal fonte de receitas advém da **filantropia** (doações individuais, doações de fundações e doações empresariais) são as que se vinculam a exercícios de avaliação mais compreensivos, estruturados, ecléticos metodologicamente e mais orientados para o impacto, quer à escala dos projetos quer à escala organizacional, dissociando-se de **exercícios de controlo e auditoria dinamizados**

por entidades financiadoras/tutelares e de práticas de reporte descritivo ao nível dos projetos, processos em que a **dependência de fundos públicos é claramente catalisadora**. É, pois, nas organizações com aquele perfil de receitas que a cultura de avaliação que encontramos maior internalização da avaliação. De resto, os **exercícios de monitorização e avaliação mais sistemáticos e formais**, sobretudo à escala organizacional e, especialmente, a mobilização de ferramentas de avaliação específicas constituem práticas que, estando fortemente associadas à priorização de informação sobre os efeitos/impactos das intervenções, **tendem a dissociar-se da adoção coerciva das práticas de avaliação e a ressaltar a sua utilidade instrumental estratégica**. Já o tipo de **abordagem quantitativa**, nomeadamente a que incide sobre realizações/*outputs*, encontra-se **vinculada às percepções mais negativas sobre a integração e implicações da prática avaliativa nas organizações**, nomeadamente à ideia de inutilidade (instrumental) e de desvio do foco essencial da intervenção, sendo portanto mais **adversas ao desenvolvimento de uma cultura avaliativa assente numa atitude coletiva de compreensão e valorização da avaliação**. Associa-se também a forças coercivas do financiamento e à procura de sobrevivência financeira. Este tipo de abordagem, que responde a imperativos de curto prazo da prestação de contas externa e se encontra desligada da missão das organizações (de mais logo prazo), tende a falhar enquanto estímulo à introdução de mudanças positivas no comportamento organizacional, conforme propus na alínea e) da segunda hipótese do estudo.

A **prevalência do financiamento público de carácter temporário** (através de projetos) na estrutura de receitas das organizações e, mais genericamente, o **envolvimento em projetos**, tende a associar-se a **abordagens quantitativas de produção de informação avaliativa**, assim como às pressões coercivas do financiamento, quando a prestação de contas se faz através de dispositivos de reporte descritivo sobre o uso dos recursos e a execução física. As avaliações formais de projetos, de natureza mais compreensiva e crítica, tendem a escapar à visão da avaliação enquanto fardo imposto pelas entidades financiadoras e a sua associação à sobrevivência organizacional, embora, tal como as outras práticas vinculadas aos projetos, se associe de forma particularmente significativa à prestação de contas ascendente. Além disso, no âmbito dos projetos, **apenas os exercícios**

de avaliação estruturados transcendem a escala dos projetos, potenciando apropriações de nível organizacional e a introdução de mudanças no comportamento individual e coletivo. O peso da influência do financiamento por projeto é, por outro lado, particularmente débil nas práticas estruturadas de avaliação e medição do desempenho/impacto de nível organizacional, e apresenta-se especialmente **desconetado da tomada de decisão e da estratégia organizacional**, estando significativamente correlacionado com usos exclusivamente não instrumentais dos exercícios avaliativos. Encontramos aqui terreno favorável aos processos de **desconexão (*decoupling*)**, nos quais as entidades se limitam a cumprir simbolicamente os preceitos mínimos de reporte, mas dissociando-os dos processos normais da organização, nomeadamente da tomada de decisão. Tal fenómeno evoca os **constrangimentos do modelo de financiamento por projeto** nas organizações do terceiro setor, nomeadamente a sobreposição dos interesses das entidades financiadoras às necessidades das organizações candidatas e o facto de o ciclo dos projetos ser muitas vezes curto e desalinhado da estratégia geral das organizações. O facto de, em **contextos de financiamento instável/temporário**, as organizações estarem particularmente vulneráveis a necessidades de informação das entidades financiadoras e a imperativos de curto prazo de prestação de contas externa desligados da missão e estratégia organizacionais, torna particularmente visíveis os **desequilíbrios entre os papéis de *accountability* e de aprendizagem organizacional**, e mais prementes as **assimetrias nas relações de poder entre entidades financiadoras e OTS**. Por outro lado, a **multiplicidade das fontes de recursos**, constituindo uma vantagem para as OTS, representa também desafios de gestão complexos para as organizações, já que encerra uma grande diversidade de regras e lógicas de controlo e avaliação que, quando impostas unilateralmente, atuam como obstáculos à internalização da avaliação.

No caso das entidades com financiamento dependente de projetos, além do padrão mais condicente com os pressupostos da teoria da agência e do isomorfismo coercivo, sobressai um arquétipo com maior proximidade aos desígnios da teoria da dependência, da teoria da *stewardship* e do isomorfismo mimético, que associa este padrão de receitas à **reciprocidade de interesses entre financiadores organizações e à possibilidade de assegurar recursos e reduzir a incerteza**. Este arquétipo encontra-se

representado no segundo perfil de avaliação revelado no estudo, no qual os mecanismos de responsabilização – entendida neste modelo numa perspetiva alargada – são essencialmente usados estrategicamente como meios para receber algo em troca (e não apenas para prestar contas), nomeadamente para atrair recursos e alavancar a obtenção de novos financiamentos. Ainda que neste modelo as contingências associadas ao financiamento por projeto e respetivos requisitos de prestação de contas ascendentes atuem como mediadores da aprendizagem, na sua vertente mais operativa, parece estabelecer-se um **equilíbrio entre os papéis de accountability e de aprendizagem organizacional**. É que, mesmo sendo a avaliação um requisito necessário, as organizações com este perfil entendem que o investimento em avaliação pode trazer algum retorno, quer em termos de eficiência quer em termos de atração de recursos.

A **participação em projetos** é, em todo o caso, fortemente impulsionadora da avaliação nas organizações estudadas, sobretudo projetos ou programas suportados por fundos públicos (nacionais ou europeus), já que ela reforça quer a **capacidade técnica** para conduzir exercícios de avaliação, quer a **intensidade da participação** na generalidade das práticas associadas à avaliação e medição do desempenho nas organizações – à exceção de alguns dos processos conotados com a gestão do desempenho (monitorização de metas organizacionais e avaliação do desempenho individual e organizacional).

Os resultados do estudo dão também suporte à **relevância das forças institucionais normativas e miméticas** decorrentes, nomeadamente, da pertença a organismos/redes internacionais. O contacto com outras organizações, potenciado pela pertença a redes de organizações (em particular, à escala internacional), assume neste estudo um lugar importante no estímulo à adoção de práticas de monitorização e avaliação nas organizações. As pressões miméticas para a adoção de práticas de medição e avaliação associadas ao contacto com entidades e profissionais de outros países decorrem especialmente do envolvimento em projetos, já que a sua influência é particularmente marcada nas práticas de monitorização e avaliação que lhes estão associadas, em especial a conceção de modelos lógicos. Nas entrevistas com OTS e outros atores relevantes sobressaem ainda as pressões normativas vinculadas ao reforço das qualificações académicas do pessoal técnico das organizações, nomeadamente ao nível da intervenção

social e da gestão, mediante a introdução de normas e padrões profissionais que norteiam, de forma crescente, o trabalho nas organizações. Sobressaem também os processos de consultoria técnica ou científica nas OTS, designadamente proveniente das universidades, mediante a introdução de práticas de diagnóstico e avaliação.

O **carácter normativo das entidades representativas do setor** para as suas associadas é, por seu turno, **pouco marcado no que toca à avaliação**, pese embora a relevância de algumas iniciativas de formação e informação dinamizadas por algumas delas. Com efeito, a pertença a organizações de cúpula não exerce um efeito significativo sobre a intensidade/abrangência da avaliação, nem na adesão a qualquer um dos exercícios de controlo, gestão e avaliação ponderados na análise. A fraca influência da pertença a estas entidades apontaria para a ideia do fraco impacto da maior estruturação do subsetor em que as OTS atuam e/ou a maior institucionalização do setor/área em que operam na internalização da avaliação. Contudo, esta ideia é, em parte, contrariada pela influência do subsetor de atividade nos padrões de adoção (e uso) da avaliação, em resultado não apenas da multiplicidade de ambientes de financiamento em que atuam ou de redes de organizações em que operam, mas também de diferentes níveis de institucionalização e autorregulação. Os **exercícios mais compreensivos e estruturados de avaliação à escala dos projetos** e, nos dois primeiros casos, mais orientados para o impacto, são característicos dos setores da **cooperação/desenvolvimento internacional**, da **participação cívica e defesa de causas** e do **desenvolvimento local e ambiente**. Este último encontra-se também particularmente adstrito ao reporte descritivo no âmbito de projetos, o qual se associa à adoção coerciva de práticas de avaliação e ao seu uso cerimonial. É também neste subsetor, a par com o da participação cívica e defesa de causas, da educação, do associativismo empresarial, profissional e sindical, e do desenvolvimento e ambiente, que a incidência da prestação de contas baseada na conformidade com os procedimentos e regulamentos é maior. Por seu turno, os exercícios de medição e avaliação do desempenho à escala organizacional, vinculados a processos de gestão estratégica e do desempenho, são genericamente menos sensíveis às forças institucionais associadas aos subsectores de atividade.

O estudo releva os **processos de adaptação institucional mimético e normativo** associados ao contexto de forte incerteza em que muitas das organizações operam, decorrentes não apenas da **instabilidade do financiamento**, mas também da crescente **concorrência entre instituições** (intensificada pela maior presença do setor mercantil no campo social), que leva as organizações a modelar-se ou a reproduzir processos e práticas de outras organizações que percebem como bem-sucedidas. A **lógica concorrencial** aproxima sobretudo as organizações das **práticas das empresas** – isomorfismo com o mercado –, impelindo-as a adotar práticas de avaliação e medição do desempenho conotadas com a gestão empresarial, designadamente no enquadramento de sistemas de avaliação do desempenho, de processos de planeamento estratégico (também particularmente vinculado a forças miméticas decorrentes da participação em plataformas internacionais) ou de sistemas de certificação da qualidade. Os processos de certificação da qualidade são identificados no estudo não apenas como fatores importantes de aposta na qualidade do serviço, mas também de diferenciação face à concorrência. Apesar de no discurso das organizações de base os processos de certificação da qualidade serem frequentemente enquadrados enquanto “processos naturais” de crescimento, profissionalização das organizações e de procura de reconhecimento e legitimidade (coerentemente com a teoria institucional), é também comum a referência a forças coercivas ligadas ao financiamento, sublimando o receio de sanções na decisão da adoção de modelos da qualidade (coerentemente com a teoria da agência). A adesão a processos de certificação da qualidade dos serviços/produtos e às práticas de avaliação e medição do desempenho que lhes estão inerentes corresponderá, pois, pelo menos num primeiro momento, mais a uma necessidade extrínseca do que intrínseca das organizações, ainda que tenda a ser normativamente internalizada como parte da boa gestão. Em todo o caso, quando falamos em monitorização e avaliação com abrangência organizacional, os **processos de gestão da qualidade**, cujas normas implicam um conjunto de novos dispositivos que acrescem aos tradicionais instrumentos de gestão, controlo, medição e avaliação, encontram-se entre os principais catalisadores, constituindo **marcos relevantes na evolução da institucionalização destas práticas nas organizações estudadas**.

Os **sistemas de gestão da qualidade**, assim como outras iniciativas de gestão mais amplas, nomeadamente os planos estratégicos, e mesmo os planos e relatórios de atividades anuais estatutariamente consagrados, têm sido **determinantes para conferir à avaliação uma estrutura mais sistematizada, formalizada e articulada com a tomada de decisão, potenciando a efetiva transferência dos ensinamentos para o ambiente organizacional**, tal como proposto na alínea d) da primeira hipótese de trabalho do estudo. A articulação dos exercícios de avaliação e medição do desempenho com estas iniciativas, ou com outros **meios de interligação dos mecanismos de geração de conhecimento com a aprendizagem organizacional é particularmente relevante no âmbito dos projetos**, no sentido de permitir uma maior integração da informação produzida a este nível nos processos de planeamento e tomada de decisão de nível organizacional. Isto é tão mais relevante quanto se percebe que o tipo de uso que se restringe à resposta às exigências dos requisitos do financiamento, sem utilidade instrumental, é particularmente marcado no contexto dos projetos, e mais ainda quando as organizações são sustentadas pelo financiamento de projetos.

Quadro 19. Visão geral das características dos tipos de avaliação

	Avaliação reativa e burocrática	Avaliação estratégica e adaptativa	
Motivação	Externa	Mista	Interna
Teorias explicativas	Teoria da agência	Teoria da dependência; Teoria da stewardship; Teoria dos stakeholders	Teoria da escolha racional; Teoria dos stakeholders
Métodos	Quantitativos	Mistos	Qualitativos
Conteúdos	Recursos e <i>outputs</i>	<i>Outputs</i> e impactos	Satisfação e impacto
Mecanismos/ Instrumentos	Relatórios execução projetos Auditorias externas	Avaliações formais projetos Ferramentas de avaliação específicas (SROI e modelos lógicos)	Monitorização e avaliação do desempenho individual e organizacional
Perfil do financiamento	Financiamento público	Filantropia/Projetos	Filantropia/receitas próprias
Prestação de contas	Ascendente	Interna, descendente e/ou ascendente	Interna e descendente
Tipo de uso	Uso simbólico	Usos concetual e instrumental	Uso instrumental

APRENDIZAGEM ORGANIZACIONAL

Fonte: Elaboração da autora.

Ora, fica claro que, tal como já tinha notado Ebrahim (2003, 2005), as organizações que fazem avaliação não sabem necessariamente como usar os resultados para guiar a mudança e alterar comportamentos, sendo para tal necessário uma ligação para a aprendizagem organizacional. Essa ligação à aprendizagem é fortemente condicionada pelas abordagens de avaliação privilegiadas nas organizações. Adotando uma perspetiva integrada sobre prestação de contas e aprendizagem organizacional, Ebrahim (2005) classifica a avaliação em três categorias muito gerais: as que fornecem uma visão dos progressos feitos num determinado período, as que alimentam a tomada de decisão organizacional e as que envolvem múltiplos constituintes. O primeiro padrão, que eu faço corresponder ao **modelo de avaliação reativo e burocrático**, identifica-se com o tipo de prestação de contas funcional, enquanto os dois últimos – que enquadro no arquétipo **estratégico e adaptativo** – se identificam com o tipo de prestação de contas estratégica, segundo a distinção avançada por Edwards e Hulme (1995) e Najam (1996) (vide quadro 18).

O **primeiro tipo de avaliação, reativo e burocrático**, só é passível de afetar as decisões e ações das OTS quando os resultados são suficientemente maus para ameaçar as fontes externas de financiamento. O seu valor é punitivo, e funciona como mecanismo de prestação de contas ascendente. Este tipo de avaliação é representado no estudo pelas abordagens quantitativas, sustentadas por dados administrativos, focadas nos recursos financeiros e nas realizações de curto prazo, como é o caso dos relatórios descritivos de execução física e financeira dos projetos, as quais, estando associada à adoção coerciva das práticas de avaliação, como vimos atrás, estão também associadas, de forma estatisticamente significativa, aos usos não instrumentais da avaliação, particularmente à prestação de contas ascendente. Donde, nos casos em que quem financia se foca em objetivos de curto prazo e demonstrações de bom emprego dos fundos, as avaliações tenderão a ser sobretudo para prestação de contas ascendente. Este modelo é prevalecte nos exercícios conduzidos no âmbito do POPH, cuja informação requerida se destina a alimentar indicadores quantificados de desempenho do programa, sem atenção a objetivos de longo prazo. O **terceiro perfil de avaliação (“avaliação como fardo e dispersão”)** identificado no ponto anterior encaixa nesta categoria.

O **modelo de avaliação estratégico e adaptativo**, correspondente ao segundo e terceiro tipos distinguidos por Ebrahim (2005), os resultados são projetados, *a priori*, para guiar as decisões das OTS, constituindo efetivamente um meio para a aprendizagem organizacional. Este modelo, que enquadra o primeiro e segundo perfis de avaliação, é representado no estudo pelos exercícios avaliativos suportados por múltiplas técnicas de produção de informação primária e secundária, qualitativa e quantitativa, e que incidem sobre os efeitos de mais longo prazo. É no seio deste modelo que encontramos terreno para a abertura ao *feedback* das diferentes partes interessadas, favorecendo a aprendizagem de longo prazo, mas também a melhoria da responsabilização descendente. Este alinhamento, particularmente sintonizado com a prestação de contas estratégica, convoca a importância da condução de processos de avaliação formal de projetos para a abertura ao envolvimento (ascendente e descendente) de atores externos. Todavia, ao contrário do que acontece ao nível da responsabilização ascendente, o papel dos projetos no reforço da responsabilização descendente é sobretudo relevante na fase de

implementação das avaliações (por via de processos de auscultação), perdendo fôlego na fase de comunicação/devolução dos resultados dos seus resultados, i.e., da prestação de contas no seu sentido mais literal. Na comunicação descendente (assim como na prestação de contas interna) sobressaem os processos de monitorização e avaliação do desempenho de nível organizacional, que tendem a enquadrar-se ou a ser vertidos em instrumentos de gestão mais amplos, como sejam os relatórios e planos de atividades, os quais são frequentemente o meio principal de comunicação entre a organização e os seus *stakeholders*. Em todo o caso, este modelo de avaliação tenderá a enquadrar relações de financiamento em que OTS e entidades tutelares/financiadoras estão mais atentas aos **objetivos de longo prazo e às dificuldades associadas à mudança social**, potenciando uma maior orientação para a aprendizagem organizacional e para a prestação de contas descendente. As relações de financiamento estabelecidas no âmbito do PIC EQUAL, são paradigmáticas deste tipo de orientação. O programa prescrevia e suportava exercícios avaliativos metodologicamente ecléticos, adaptativos, focados na mudança e norteados por valores e princípios, um dos quais a participação de todas as partes interessadas na intervenção (ainda que esta recomendação do programa tenha sido debilmente acolhida pelos projetos).

Muito embora o trabalho enfatize o papel significativo dos fatores estruturais e ambientais, **os resultados não afastam a importância do contexto organizacional interno das organizações** na compreensão dos padrões de adoção e uso da avaliação. O estudo sublinha a relevância de algumas variáveis associadas à **cultura organizacional**, enquanto atributos internos das organizações, para a internalização da avaliação. Fica claro que o contexto interno associado a perfis organizacionais com menor adesão à avaliação tende a ser marcado por uma cultura organizacional hostil à avaliação, assente quer numa liderança pouco aberta e/ou comprometida com processos de reflexão, análise e aprendizagem contínua, quer numa atitude geral reativa à avaliação por parte dos membros da avaliação. A atitude individual e coletiva de compreensão, aceitação e valorização da avaliação que caracteriza o contexto interno dos perfis organizacionais com maior adesão à avaliação é, por seu turno, tributária de um clima de abertura, confiança e encorajamento suportado no apoio e/ou envolvimento das lideranças e da perceção de

utilidade e/ou retorno externo destes exercícios. Estes resultados sublinham a **relevância da agência dos *stakeholders* internos** (em especial dos líderes das organizações) no que respeita à implementação de processos de avaliação, sugerindo algum nível de controlo dos membros da organização sobre as decisões associadas à avaliação e, portanto, ultrapassando as visões deterministas sobre a influência de fatores ambientais (Barman e MacIndoe, 2012).

CONSIDERAÇÕES FINAIS

Articulando um conjunto amplo de perspectivas teóricas e analíticas, radicadas em campos do conhecimento diversos, este trabalho fornece contributos adicionais para caracterizar as práticas de avaliação nas OTS e para compreender os fatores e padrões associados à sua adoção e ao seu uso, concorrendo para colmatar lacunas teóricas e empíricas neste domínio.

O estudo contribui para alimentar o entendimento da avaliação como conceito de múltiplas significações, que se encontram refletidas na **amplitude de práticas e lógicas avaliativas** e na diversidade de enquadramentos das atividades de avaliação nas organizações do terceiro setor, **permeáveis que são às racionalidades de avaliação produzidas nos polos que intercetam o setor - Estado, mercado e comunidade**. Situando as lógicas de avaliação das organizações do terceiro setor num **campo híbrido entre diversos tipos de discursos e práticas** associados a cada um desses três polos, o estudo revelou elementos que apontam para fenómenos de complementaridade, mas também de colonização de racionalidades. A prevalência do tipo de perspectiva tecnocrática e managerialista a partir da qual a avaliação é concebida, em termos do seu conteúdo e desenho metodológico, e a secundarização da linguagem que reflete valores associativos face à linguagem alicerçada em outros domínios discursivos (sobretudo a originada na cultura gestionária), parecem apontar para a subordinação do caráter “expressivo” das organizações (isto é, de expressão de valores e ideais de serviço), ao seu caráter “instrumental” (isto é, à sua orientação material, racional e objetiva).

Com efeito, a avaliação tende a assumir no estudo feições próximas das práticas de **controlo de gestão do setor empresarial**, em que as análises mais compreensivas e contextuais cedem lugar a uma lógica mais alinhada com os pressupostos da medição do desempenho, e que pode ser enquadrada na primeira e segunda gerações da avaliação, que Guba e Lincoln (1989) nomeiam de fases da **“medição” e da “descrição”**. A lógica de avaliação prevalecente nas OTS estudadas, focada no controlo, na monitorização interna e na conformidade com métricas, padrões e objetivos predeterminada reflete, pelo menos em parte, os requisitos associados ao financiamento (nomeadamente de reporte descritivo no âmbito de projetos e de acompanhamento externo/auditoria), assim como disposições

dos sistemas de gestão estratégica e do desempenho e de certificação/acreditação do TS, eminentemente orientados para a relação entre *inputs/outputs*, eficiência/eficácia, e fortemente arreigados na filosofia de gestão por objetivos.

Não obstante a simplicidade e “produtividade” deste tipo de abordagem dominante nas OTS, com ela as organizações relegam **o potencial da pesquisa em avaliação para a análise e compreensão mais aprofundadas** de como e porquê uma intervenção (não) funciona ou determinadas mudanças ocorrem. São negligenciadas, por outro lado, as abordagens “de baixo para cima”, que se abrem à realidade subjetiva (isto é, às diversas perspetivas das partes interessadas), prevalecendo uma perspetiva de eficácia, enquanto medida de “valor” ou “sucesso”, essencialmente assente na capacidade para cumprir objetivos. Ao fazer assentar a avaliação no cumprimento de objetivos – sobretudo quando suportados por indicadores quantitativos – negligenciam-se perspetivas particularmente relevantes em entidades com uma missão social e que atuam em ambientes complexos e incertos, como são as OTS, ao ser omitida informação de contexto e sobre os efeitos não previstos ou fora do âmbito dos objetivos fixados pelas organizações¹⁸⁸. **Ocultam-se ainda dimensões de processo particularmente relevantes para aferir sobre o desempenho das OTS**, como sejam a participação, o pluralismo, a equidade, a advocacia e inovação (Kendall e Knapp, 2000). Isto porque a perspetiva de *processos* que prevalece nas organizações estudadas centra-se sobretudo em critérios de economia, eficiência interna e eficácia “intermédia”, isto é, na capacidade de fazer uma gestão eficiente dos recursos e de cumprir os *outputs* propostos (em termos de atividades e pessoas abrangidas), e não em critérios como o envolvimento e mobilização dos cidadãos e cidadãs ou a inovação na implementação das intervenções. Ora, é **exatamente este tipo de processo que compõe o carácter distintivo do setor e que, por isso, deverá integrar qualquer esforço para avaliar a sua eficácia.**

¹⁸⁸ Estas limitações sugerem que, como diversos trabalhos têm proposto (e.g., Ebrahim, 2003a; D’Aunno, 1992), embora faça sentido conduzir avaliações a fim de apreciar o progresso em relação aos objetivos, este não deve ser o único, ou mesmo o propósito essencial da avaliação. Se nalguns contextos se adequará o recurso ao modelo de avaliação por objetivos, por exemplo, quando o trabalho da organização é fácil de medir e quando os critérios de desempenho tendem a ser partilhados por diversos grupos de *stakeholders*. Já em situações em que os critérios de desempenho variam entre *stakeholders* ou em que a missão organizacional se orienta para processos (e.g., advocacia política ou *empowerment*), uma abordagem reputacional/de múltiplos constituintes será mais adequada.

A ausência da perspectiva de avaliação ancorada nos efeitos/mudanças (nas pessoas beneficiárias ou na sociedade) nas organizações estudadas constitui-se simultaneamente como causa e efeito do débil envolvimento dos públicos-alvo e outros *stakeholders* nos processos avaliativos, compondo um modelo de prestação de contas eminentemente funcional. Apesar dos esforços já visíveis de envolvimento das partes interessadas, o estudo mostra que o **potencial da avaliação para o fortalecimento da responsabilização descendente através de processos participativos** (Edwards e Hulme, 1996; Benjamin, 2013) ainda se encontra **largamente inexplorado**. Encontra-se também pouco capitalizado o potencial dos processos de avaliação pluralistas, colaborativos, para o aprofundamento da inclusão democrática e emancipação (Cousins e Whitmore, 1998) e para a aprendizagem de nível conceptual (Chouinard, 2013).

A abordagem managerialista de **medição do desempenho** prevalecente nas OTS demonstra, porém, **não ser incompatível** com domínios relevantes de **aprendizagem organizacional**, nomeadamente os associados ao uso instrumental/material dos resultados. A integração da avaliação em sistemas de gestão mais amplos, originados e popularizados no setor empresarial, como é o caso dos sistemas de gestão da qualidade e de gestão estratégica (nomeadamente o planeamento estratégico), tem-se revelado proveitosa não só para conferir à avaliação uma estrutura mais formalizada, mas também mais articulada com os processos de decisão de nível organizacional, reforçando o seu funcionamento efetivo para a melhoria do desempenho e, em alguns casos, até mais alinhada com a missão organizacional. Será, pois, proveitoso potenciar a articulação destes sistemas com a avaliação. Não fazer a distinção entre estes exercícios é, contudo, potencialmente problemático, pois mitiga a relação funcional entre a avaliação e aquelas ferramentas de gestão, já que a avaliação é necessária para informar a gestão de qualidade e o planeamento estratégico (e operacional) e os resultados da avaliação devem ser usados para melhorar o desempenho (Carman e Fredericks, 2008; Ellis e Gregory, 2008). Além disso, ao mesmo tempo que a integração da avaliação em estruturas de gestão mais abrangentes canaliza o seu funcionamento para a melhoria do desempenho, o seu papel político e de aprendizagem mais amplo poderá ser desvalorizado.

Por outro lado, sendo claras as manifestações de isomorfismo com o Estado e com o mercado no que toca à avaliação, mediante a prevalência de técnicas de gestão racionais e padronizadas, e alguma diluição dos valores que definem as contribuições distintivas do terceiro setor, também são encontrados **sinais de combinação e até reforço mútuo** entre as racionalidades emanadas do Estado e do mercado, e as originadas na sociedade civil. Vemos que em algumas organizações com processos de gestão profissionalizados – nomeadamente a partir de sistemas de certificação da qualidade concebidos no seio do TS e para o TS (e.g., EQUASS) – ou com processos de avaliação concebidos e implementados profissionalmente – especialmente os exercícios mais estruturados, compreensivos e participativos de avaliação à escala dos projetos –, a forte incidência de práticas e de linguagem com raízes na cultura gestionária (e/ou científica) articula-se com o maior uso de métodos e linguagem que refletem valores associativos. Nestas combinações, a avaliação é vista como tendo não só **propósito instrumental** de melhorar o desempenho mas também uma **qualidade expressiva**, assente na preservação dos valores constitutivos das organizações e pessoas envolvidas. O estudo mostra que as organizações marcadamente híbridas, que acumulam funções de serviço com a expressão de valores, interesses e crenças (culturais, espirituais, profissionais ou políticas), são exatamente aquelas que, misturando diferentes racionalidades, tarefas, e papéis diferentes, mais evidenciam estas combinações e mais delas parecem beneficiar para a eficácia organizacional.

De resto, vimos que, em termos mais gerais, a **combinação das diversas lógicas de avaliação** (científica, gestionária e associativa) no discurso das organizações é sobretudo devedora da **articulação das preocupações técnicas**, em torno do desempenho e dos resultados, com a **centralidade do objetivo de serviço à comunidade ou a grupos específicos de pessoas**. Todavia, apesar da centralidade dos públicos-alvo e da comunidade no discurso das organizações, o estudo sinaliza o enviesamento dos processos de avaliação em favor dos **atores “ascendentes”**, uma vez que a capacidade de exigir prestação de contas é claramente mais forte por parte das entidades financiadoras do que por parte de quem as OTS procuram servir. Ora, sendo a população-alvo das organizações a marca mais visível da génese associativa da linguagem partilhada pelas organizações, a consideração

dos seus interesses e expectativas deverá ser central no enquadramento das estratégias de comunicação e coordenação através das fronteiras sociais e linguísticas.

Ao explorar o modo como diversas teorias organizacionais podem ser mobilizadas para compreender as práticas de avaliação, o estudo contribui para **questionar os pressupostos de racionalidade** e de causalidade que têm regido muita da discussão teórica e metodológica sobre avaliação, sublinhando a sua natureza relacional (Ebrahim, 2003a, 2003b; Beis, 2010; Chouinard, 2013) e o seu enquadramento enquanto mecanismo de prestação de contas/responsabilização. A análise evidencia que as pressões coercivas associadas ao financiamento, as normativas associadas à certificação, ou as miméticas associadas à pertença a redes de organizações (nomeadamente à escala internacional), assumem um lugar importante no estímulo à adoção de práticas de monitorização e avaliação nas organizações, embora produzam efeitos diferenciados na internalização da cultura avaliativa.

Ao desvelar a composição de modelos/perfis de avaliação distintos e com implicações diversas na aprendizagem organizacional, o estudo contribui para **alimentar a discussão política sobre o papel da avaliação no reforço da responsabilização**, fornecendo pistas importantes para a (re)definição dos processos de prestação de contas, nomeadamente no âmbito das relações de financiamento. Desde logo, a multiplicidade de partes interessadas das OTS (que não se esgotam nas relações de financiamento) reivindica a adoção de uma perspetiva ampla de prestação de contas que não subordine os mecanismos e atores descendentes e internos aos ascendentes. A análise chama a atenção para a ineficácia das relações de *accountability* ascendente de carácter coercivo e punitivo, assentes na conformidade com requisitos de descrição e quantificação que respondem às agendas de quem financia, enquanto reivindica as mais-valias do estabelecimento de relações de *accountability* baseadas na *stewardship*, assentes em modelos de avaliação mais abertos e adaptativos, e mais vinculadas aos valores e propósitos de longo prazo das organizações. Embora o uso da avaliação para a prestação de contas desempenhe um papel legítimo e necessário, este deverá ser entendido apenas como um dos propósitos da avaliação. Tal como os dados do estudo indicam, quando as abordagens à avaliação nas organizações não se esgotam nesta perspetiva, e quando à função concetual e simbólica

da avaliação se associa uma função utilitária, os ganhos são mais significativos, quer em termos de desempenho quer no reforço da relação com entidades financiadoras e outros *stakeholders*. Será por isso essencial sinalizar a importância da avaliação, nas suas diversas funcionalidades (instrumentais e não instrumentais), desenvolvendo abordagens consequentes e efetivamente compatíveis com os benefícios potenciais da avaliação, e **valorizando o investimento que é feito nestes processos**. Tal implica **gerir uma ambivalência** detetada no estudo na forma como a avaliação é acolhida: se, por um lado, a perceção de que o investimento em avaliação não é valorizado externamente é dissuasora deste tipo de investimento, por outro lado, a perspectiva de sanções associada à condução de exercícios de avaliação é desmobilizadora para as pessoas que neles participam e abre caminho a fenómenos de manipulação ou distorção da informação avaliativa. Uma forma de contornar esta dualidade por parte de entidades financiadoras e outros *stakeholders* será assumir uma perspectiva que, mais do que centrar-se na demonstração de resultados por parte das OTS, i.e., demonstrar que estão a “fazer um bom trabalho”, se centre na sua capacidade de usar os processos e resultados da avaliação de forma significativa para a prossecução da missão, i.e., demonstrar que estão a fazer avaliação “para fazer um trabalho melhor”. A análise reivindica também a necessidade de reorientar os requisitos de avaliação para a recolha de informação qualitativa sobre uma maior amplitude de dimensões associadas aos processos e efeitos da intervenção das organizações, afastando-se do modelo de relatório focado em números de desempenho, irrelevante para a aprendizagem organizacional.

De resto, acompanhando as conclusões deste estudo, têm sido várias as análises a dar conta dos **riscos** associados à adoção de **abordagens de avaliação sumativas que premeiam o sucesso e punem as falhas** (e.g., suspendendo ou condicionando o financiamento), particularmente mediante **modelos de medição quantitativos** (Ebrahim 2003a; Barman, 2007; Salamon *et al.*, 2010; Benjamim, 2012; Edwards *et al.*, 2014; Ferreira, 2017). O perigo de desvio da missão é crível, já que este tipo de abordagem tenderá a orientar a ação das OTS para intervenções e para grupos sociais que possuem menores riscos de fracasso, dissuadindo abordagens inovadoras e audaciosas. Tenderá também a direcionar as organizações para a consecução de resultados mais fáceis de medir,

desconsiderando efeitos difíceis de quantificar (*soft outcomes*), como são os que se vinculam às funções expressivas das organizações (e.g., papéis de militância e de promoção da participação). A consideração destes riscos assume no atual contexto português uma relevância acrescida, em face da afirmação de novas lógicas de financiamento ao setor, as quais integram este tipo de abordagem. É o caso da contratualização de resultados (pagamento por resultados) e da disponibilização de instrumentos financeiros distantes da lógica tradicional de donativo, como sejam o investimento social/de impacto e os Título de Impacto Social, presentes no novo ciclo de programação dos fundos europeus 2014 -2020 (Portugal 2020).

O impacto do tipo de abordagem de avaliação privilegiada pelas entidades financiadoras não é uniforme nas OTS, já que elas constituem um grupo heterogéneo, com propósitos, missões, valores e papéis diversos. O olhar do estudo sobre o fatores de contingência associados à identidade e estrutura organizacionais, como sejam o tipo de entidade, a dimensão, a idade e funções organizacionais, vem reforçar a tese das **implicações diferenciadas da diversidade de vocações, atividades e capacidades na configuração da avaliação das organizações do terceiro setor**, sinalizando a necessidade de desenvolver abordagens sensíveis e ajustadas às suas especificidades.

O estudo releva, por exemplo, o **desafio acrescido** que se coloca às organizações de **advocacia ou expressão de interesses** – cujas mudanças produzidas têm um caráter particularmente intangível e dificilmente atribuível à sua ação – **para medir os efeitos da sua ação**, sobretudo quando estruturada a partir de ferramentas de avaliação ou medição do desempenho padronizadas, como é o caso do SROI e dos “modelos lógicos”. Para estas organizações, as abordagens de avaliação mais acessíveis terão, pois, um caráter mais aberto e menos padronizado e com foco em resultados intermédios - que possam ser mais diretamente atribuíveis à sua intervenção ou sobre os quais possuam maior controlo (Ebrahim e Rangan, 2010, 2014).

No que toca à dimensão, o estudo é revelador dos **constrangimentos acrescidos que se colocam às organizações de pequena dimensão**, sobretudo em termos de recursos e infraestruturas para avaliação. As abordagens qualitativas de produção de informação de avaliação, sendo menos exigentes em termos de custos e competências, serão mais

acessíveis às organizações de pequena dimensão financeira, nos quais o financiamento para a avaliação é limitado ou inexistente. Algumas das ferramentas/técnicas qualitativas de avaliação disponíveis (e.g., Planeamento de Impacto, Avaliação e Aprendizagem - IPAL) foram, exatamente, desenvolvidas nas e para organizações do terceiro setor, e por isso mais cuidadosamente harmonizadas com as exigências que as técnicas de avaliação podem colocar em termos de recursos e competências (e.g., Hall, 2014; Davies, 1998), quando comparadas com as abordagens quantitativas desenvolvidas por entidades financiadoras e doadoras (como é o caso do SROI e do quadro lógico).

Por fim, a análise é reveladora da **“fragilidade”** das **organizações de constituição mais antiga**, de alguma forma já legitimadas, e portanto menos predispostas à introdução de práticas de avaliação que sinalizem o valor da sua intervenção, enquanto sublinha a maior abertura das organizações de constituição mais recente ao investimento em processos de monitorização e avaliação, nomeadamente os que envolvem a recolha de informação sobre as mudanças, ainda que o enquadramento dos processos de avaliação em iniciativas de gestão mais amplas, que conferem à avaliação uma estrutura mais formalizada, como a gestão estratégica ou da qualidade, exija uma maturidade organizacional ainda não consolidada nas organizações mais jovens. As organizações de constituição mais recente, por serem mais orientadas para a contratação de pessoal especializado nas áreas de suporte ao desenvolvimento das atividades (Hwang e Powell, 2009), mais conscientes das expectativas do campo organizacional e menos influenciadas pelos “hábitos” organizacionais (Barman e MacIndoe, 2012), tenderão a ser mais recetivas à adoção de práticas racionalizadas, particularmente as radicadas no mercado (assentes em lógicas de medição).

O carácter heterogéneo do setor desafia, pois, a elaboração de uma matriz analítica única, e convoca o desenvolvimento de abordagens diferenciadas e multidimensionais que acomodem a sua complexidade. Esta diversidade compõe a vantagem destas organizações e a mais-valia do seu contributo para a sociedade, mas **é esta mesma heterogeneidade que questiona a utilidade de abordagens “one-size-fits-all”, que desconsideram a singularidade das suas funções e papéis.**

A abordagem metodológica mista em que assentou esta investigação permitiu analisar os contextos, os discursos e as práticas de avaliação nas OTS a partir de uma diversidade de ângulos. Não obstante, o desenho da pesquisa revelou algumas **limitações**, a partir das quais é possível identificar **áreas para pesquisa futura**. Em primeiro lugar, a seleção das amostras de OTS a observar teve por base a decisão de acolher a composição heterogénea do terceiro setor – integrando as suas diferentes famílias, em termos de formas e estatutos jurídicos, perfis de atividade, dimensões, idades e localizações geográficas. Esta opção estendeu a amplitude de situações em termos da integração de práticas de avaliação, sendo incluído no estudo um grande número organizações sem práticas de avaliação consistentes e estruturadas. A subamostra de organizações com práticas de avaliação significativas e sistematizadas é, por isso, relativamente pequena, o que poderá ter condicionado a capacidade do estudo para compreender a extensão dos contextos e condições associados a uma maior internalização da cultura avaliativa. Em segundo lugar, a produção de informação através do inquérito às OTS assentou essencialmente na visão de uma pessoa (geralmente com funções de liderança), que atuou em representação de toda a experiência da organização, o que restringiu a possibilidade de compreender a multiplicidade de sentidos, perspetivas e perceções sobre avaliação no seio de cada uma das organizações. A realização de estudos complementares, nomeadamente de carácter etnográfico, que envolvam uma maior diversidade de *stakeholders* (internos e externos) das organizações, poderá fornecer uma imagem mais completa, detalhada e matizada do estado da avaliação nas OTS. Em terceiro lugar, e nesta mesma linha, seria útil estudar em maior profundidade algumas dinâmicas (potencialmente) influentes da avaliação associadas ao contexto interno das organizações, nomeadamente as relativas à cultura organizacional (e.g., adaptação, aprendizagem, inovação), mas também as alusivas a características de governação e gestão das organizações (e.g., modelos de governação e capacidade de governo, tipos de liderança, modos de gestão). Por fim, seria proveitoso que, em pesquisas futuras, fosse testada, através de métodos estatísticos, a relevância da relação da adesão à avaliação e suas dinâmicas de implementação com um conjunto mais amplo de variáveis associadas ao contexto institucional, como sejam a perceção das organizações sobre as práticas

avaliativas de outras entidades do campo organizacional (DiMaggio e Powell, 1983; Eckerd e Moulton, 2011; Barman e MacIndoe, 2012) e a exposição a lógicas institucionais do mercado (Barman e MacIndoe, 2012).

Apesar das limitações apontadas, o estudo dá-nos um contributo extremamente importante para o enquadramento, contextualização e compreensão das práticas de avaliação nas OTS portuguesas, que se espera vir a ser útil, nomeadamente, para alimentar a discussão política e metodológica – ou a “política do método”, como a designa Greene (2009) – no campo da avaliação e da prestação de contas.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alaimo, Salvatore P. (2008) "Nonprofits and evaluation: Managing expectations from the leader's perspective" in Carman, Joanne; Fredericks, Kimberly (eds.) *Nonprofits and evaluation*. New Directions for Evaluation, 119, 73–92.
- Alkin, Marvin C. (2013) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: Sage. (2.ª Edição)
- Alkin, Marvin C.; Taut, Sandy (2003) "Unbundling Evaluation Use" *Studies in Educational Evaluation*. 29 (3), 1–12.
- Almeida, Vasco (2011) *As Instituições Particulares de Solidariedade Social. Governação e terceiro setor*. Coimbra: Editora Almedina.
- Anheier, Helmut K. (2005) *The Nonprofit Setor: Approaches, Management, Policy*. London e New York: Routledge.
- Arena, Marika; Azzone, Giovanni; Bengo, Irene (2015) "Performance Measurement for Social Enterprises" *Voluntas*. 26 (2), 649–72.
- Arvidson, Malin (2009), Impact and Evaluation in the UK Third Sector: Reviewing Literature and Exploring Ideas. Third Sector Research Centre *Working Paper 27*, Birmingham: TSRC, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-27.pdf> [18 de Abril de 2012].
- Arvidson, Malin; Kara, Helen (2013) "Putting Evaluations to Use : From Measuring to Endorsing Social Value, *TSRC Working Paper 110*, Birmingham: Third Sector Research Center, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-110.pdf> [19 de Setembro de 2014]
- Arvidson, Malin; Lyon, Fergus (2014) "Social Impact Measurement and Non-Profit Organisations: Compliance, Resistance, and Promotion" *Voluntas*, 25, 869–86.
- Arvidson, Malin; Lyon, Fergus; McKay, Stephen; Moro, Domenico (2013) "Valuing the Social? The Nature and Controversies of Measuring Social Return on Investment (SROI)" *Voluntary Sector Review*. 4 (1), 3–18.
- Arvidson, Malin; Lyon, Fergus; McKay, Stephen; Moro, Domenico (2010) "Ambitions and Challenges of SROI" *Third Sector Research Centre Working Paper 49*, Birmingham: TSRC, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-49.pdf> [22 de Setembro de 2013].
- Bagnoli, Luca; Cecilia, Megali (2011) "Measuring Performance in Social Enterprises" *Nonprofit and Voluntary Sector Quarterly*. 40, 149–65.

- Ballart, Xavier (1992) *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudios de caso*. Madrid: Ministerio para las Administraciones Públicas.
- Barman, Emily (2007) "What Is the Bottom Line for Nonprofit Organizations? A History of Measurement in the British Voluntary Sector" *Voluntas* 18 (2), 101–15.
- Barman Emily; MacIndoe, Heather (2012) "Institutional Pressures and Organizational Capacity: The Case of Out- come Measurement" *Sociological Forum*. 27 (1): 70 – 93.
- Baruch, Yehuda; Nelson, Ramalho (2006) "Communalities and Distinctions in the Measurement of Organizational Performance and Effectiveness Across For-Profit and Nonprofit Sectors". *Nonprofit and Voluntary Sector Quarterly*, 2006, 35(1), 39–65.
- Bastoe, Per O. (2006) "Implementing Results-Based Management" in Stame, Nicoletta; Rist, Ray (eds.) *From Studies to Streams: Managing Evaluative Systems*. London: Transaction Publishers, 97–110.
- Benjamin, Lehn M. (2008) "Account Space: How Accountability Requirements Shape Nonprofit Practice" *Nonprofit and Voluntary Sector Quarterly*. 37 (2), 201–23.
- Benjamin, Lehn M. (2012) "Nonprofit Organizations and Outcome Measurement: From Tracking Program Activities to Focusing on Frontline Work" *American Journal of Evaluation*. 33 (3), 431–47.
- Benjamin, Lehn M. (2013) "The Potential of Outcome Measurement for Strengthening Nonprofits' Accountability to Beneficiaries" *Nonprofit and Voluntary Sector Quarterly*. 42 (6), 1224–44.
- Bernstein, David J. (1999) "Comments on Perrin's 'Effective use and misuse of performance measurement'" *American Journal of Evaluation*. 20, 85–93.
- Bickel, William E.; Millet, Ricardo; Nelson, Catherine (2003) *Challenges to the role of evaluation in supporting organizational learning in foundations*. Washington, DC: Council on Foundations.
- Bies, Angela, L. (2010) "Evolution of Nonprofit Self-Regulation in Europe" *Nonprofit and Voluntary Sector Quarterly*. 39 (6), 1057–1086.
- Bjørnholt, Bente; Larsen, Flemming (2014) "The Politics of Performance Measurement: 'Evaluation Use as Mediator for Politics" *Evaluation*. 20 (4), 400–411.
- Boswell, Christina (2008) "The political functions of expert knowledge: knowledge and legitimation in European Union immigration policy" *Journal of European Public Policy*. 15, 471–88.
- Botcheva, Luba; White, Catherine R.; Huffman, Lynne C. (2002) "Learning culture and outcomes measurement practices in community agencies" *American Journal of Evaluation*, 23 (4), 421–434.

- Bourgeois, Isabelle, Cousins, J. Bradley (2008) "Informing Evaluation Capacity Building through Profiling Organizational Capacity for Evaluation: An Empirical Examination of Four Canadian Federal Government Organizations" *Canadian Journal of Program Evaluation*. 23 (3), 127–46.
- Bourgeois, Isabelle; Cousins, J. Bradley (2013) "Understanding dimensions of organizational evaluation capacity" *American Journal of Evaluation*. 34(3), 299–319.
- Boyle, Richard; Lemaire, Donald, & Rist, Ray (1999) "Introduction: Building evaluation capacity" In Boyle, Richard; Lemaire, Donald (eds.) *Building effective evaluation capacity: Lessons from practice*. New Brunswick, NJ: Transaction, 1-22.
- Bradford, Neil, Chouinard, Jill A. (2009) "Learning through evaluation? Reflections on two federal community-building initiatives" *Canadian Journal of Program Evaluation*. 24(1), 51–77.
- Breckell, Paul; Harrison, Kate; Robert, Nicola (2010) *Impact Reporting in the UK Charity Sector*. London: Charity Finance Group/ Cass Business School. <http://www.cfg.org.uk/resources/~media/Files/Resources/Impact%20Reporting%20in%20the%20UK%20Charity%20Sector.ashx> [11 de Março de 2013].
- Brisolara, Sharon; Seigart, Denise; SenGupta, Saumitra (eds.) (2014) *Feminist evaluation and research: Theory and practice*. New York: Guilford Press.
- Brown, Gillian; Yule, George (1983) *Discourse Analysis*. Linguistics Cambridge: Cambridge University Press.
- Brunsson Nils (2002) *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*. Copenhagen: Abstrakt & Liberc.
- Bryman, Alan (2012) *Social Research Methods*. New York: Oxford University Press.
- Bryman, Alan; Becker, Saul; Sempik, Joe (2008) "Quality Criteria for Quantitative, Qualitative and Mixed Methods Research: The View from Social Policy" *International Journal of Social Research Methodology*. 11, 261– 76.
- Bryson, John M. (2011) *Strategic planning for public and nonprofit organizations*. San Francisco: Jossey-Bass. (4.ª Edição).
- Cameron, Kim S., Whetton, David A. (1983) *Organisational effectiveness: a comparison of multiple models*, New York: Academic Press.
- Campbell, David (2002) Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management and Leadership*, 12(3), 243-259.
- Cano, Ignacio (2006) *Introdução a avaliação de programas sociais*. Rio de Janeiro: FGV.
- Capucha, Luís (2008) *Planeamento e Avaliação de Projetos - Guia Prático*. Lisboa: DGIDC.

- Capucha, Luís; Almeida, João; Pedroso, Paulo; Silva, José (1996) "Metodologias de avaliação: o estado da arte em Portugal" *Sociologia - Problemas e Práticas*, 22, 11-17.
- Carey, Gemma; Braunack-Mayer, Amnette; Barraket, Jo (2009) "Spaces of care in the third sector: Understanding the effects of professionalization" *Health*. 13, 629-646.
- Carman, Joanne G. (2007) "Evaluation Practice Among Community-Based Organizations: Research Into the Reality" *American Journal of Evaluation*. 28 (1), 60–75.
- Carman, Joanne G. (2010) "The accountability movement: what's wrong with this theory of change?" *Nonprofit and Voluntary Sector Quarterly*. 39 (2), 256 -274
- Carman, Joanne G. (2011) "Understanding Evaluation in Nonprofit Organizations" *Public Performance & Management Review*. 34 (3), 350–77.
- Carman, Joanne G.; Fredericks, Kimberly A. (2008) "Nonprofits and evaluation: Empirical evidence from the field", *Nonprofits and evaluation. New Directions for Evaluation*. 51–71.
- Carman, Joanne G.; Fredericks, Kimberly A. (2009) "Evaluation Capacity and Nonprofit Organizations: Is the Glass Half-Empty or Half-Full?" *American Journal of Evaluation*. 31 (1), 84-104.
- Carmel, Emma; Harlock, Jenny E. (2008) "Instituting the 'Third Sector' as a Governable Terrain: Partnership, Procurement and Performance in the UK" *Policy and Politics*. 36 (2), 155–71.
- Carvalho, Ana (2010) "Quantifying the third setor in Portugal: An overview and evolution from 1997 to 2007" *Voluntas*. 21, 588–610.
- Chambers, Robert (1994) *Challenging the Professions*. London: Intermediate Technology Publications.
- Chapman, Tony; Robinson, Fred; Brown, Judith; Crow, Robert; Bell, Victoria; Bailey, Emma (2010) *What Makes a Third Sector Organisation Tick? Interactions of Foresight , Enterprise , Capability and Impact*. Northern Rock Foundation TSTS. <http://www.nrfoundation.org.uk/downloads/NRF-TST-Report-What-makes-TSO-tick.pdf> [11 de Março de 2013].
- Chouinard, Jill (2013) "The Case for Participatory Evaluation in an Era of Accountability" *American Journal of Evaluation*. 34 (2), 261–65.
- Christie, Cristina; Alkin, Marvin (2013) "An Evaluation Theory Tree" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: Sage, 11-57. (2.ª Edição).
- Chung, Phillip; Tran, Melanie (2015) *State of Evaluation in Colorado's Nonprofit Sector*. Denver CO: The Colorado Trust.

- <http://www.coloradotrust.org/sites/default/files/state%20of%20evaluation%208-2015%20links.pdf> [28 de Outubro de 2015].
- Cohen, Ernesto; Franco, Rolando (2006) *Avaliação de projetos sociais*. Petrópolis: Vozes. (6.ª Edição)
- Cohen, Rick (2007) "Conflicts of interest: Mischief, thou art afoot" *Nonprofit Quarterl.* 14(1), 54-60.
- Comissão Europeia (CE) (2003) *EVALSED: A Avaliação do Desenvolvimento Socioeconómico*. Directorate-General Regional Policy. Brussels: CE.
- Comissão Europeia (CE) (2012) *Proposed Approaches to Social Impact Measurement in European Commission legislation and in practice relating to EuSEFs and the EaSI*. Luxembourg: Publications Office of the European Union.
- Comissão Europeia (CE) (2013) *EVALSED: The resource for the evaluation of Socio-Economic Development*. Directorate-General for Regional and Urban Policy, Brussels: CE, http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf [22 de Abril de 2014]
- Comité Económico e Social Europeu (CESE) (2007) *A Economia Social na União Europeia*. Brussels: CESE, <http://www.eesc.europa.eu/resources/docs/eesc-2007-11-pt.pdf> [8 de Outubro de 2015].
- Connolly, Terry; Conlon, Edward J; Deutsch, Stuart J. (1980) "Organizational effectiveness: A multiple- constituency approach" *The Academy of Management Review.* 5, 211–217.
- Cornwall, Andrea; Lucas, Henry; Pasteur, Kath (2000) "Introduction: Accountability through participation: Developing workable partnership models in the health sector" *IDS Bulletin*, 31(1), 1-13.
- Cousins, J Bradley; Goh, Swee C.; Clark, Shannon; Lee, Linda E. (2004) "Integrating Evaluative Inquiry into the Organizational Culture: A Review and Synthesis of the Knowledge Base" *New Directions for Evaluation.* 19 (2), 99–141.
- Cousins, J. Bradley (2013) "Understanding Use Through Systematic Inquiry" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, 344-351. (2.ª Edição)
- Cousins, J. Bradley; Earl, Lorna M. (1992) "The Case for Participatory Evaluation" *Educational Evaluation and Policy Analysis*, 21 (14), 397-418.
- Cousins, J. Bradley; Whitmore, Elizabeth (1998) "Framing participatory evaluation" *New Directions for Evaluation*, 80, 5–23.

- Creswell, John (1998) *Qualitative Inquiry and Research Design: Choosing among five traditions*. Thousand Oaks: SAGE Publications.
- Cruz, Sofia; Pais, Celso; Parente, Cristina (2014) “Gestão Estratégica, Liderança e Cultura nas Organizações do Terceiro Setor” in Parente, Cristina (ed.) *Empreendedorismo Social em Portugal*. Porto: Universidade do Porto.
- Cuesta, Marisa; Martínez, Ana (2010) *Fundamentos de Evaluación de Políticas Públicas*. Ministerio de Política Territorial Y Administración Pública Agencia Estatal de Evaluación de Las Políticas Públicas Y La Calidad de Los Servicios. doi:10.1073/pnas.0703993104.
- Cutt, James; Murray, Vic (2000) *Accountability and Effectiveness Evaluation in Non-Profit Organizations*. Abingdon: Taylor & Francis.
- Davis, James; Schoorman, F. David; Donaldson, Lex (1997) “Toward a Stewardship Theory of Management” *Academy of Management Review*. 22 (1), 20-47.
- Dees, J. Gregory (1998) “Enterprising nonprofits” *Harvard Business Review*, 76(1), 54-66.
- Dees, J. Gregory; Anderson, Beth B. (2006) “Framing a Theory of Social Entrepreneurship: Building on Two Schools of Practice and Thought” *Research on Social Entrepreneurship, ARNOVA Occasional Papers Series*.
- Development Assistance Committee (DAC) (2000) *Results Based Management in the Development Co-Operation Agencies: A Review of Experience*. Paris: OCDE.
- Development Assistance Committee (DAC) (2010) *Glossary of Terms in Evaluation and Results-Based Management*. Paris: OECD.
- DiMaggio, Paul (2001) “Measuring the Impact of the Nonprofit Sector on Society is Probably Impossible but Possibly Useful” in Patrice Flynn; Hodgkins, Virgínia A. (Eds.), *Measuring the impact of the nonprofit sector*. New York: Kluwer Academic, 249-272.
- DiMaggio, Paul; Powell, Walter (1983) “The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields” *American Sociological Review*. 48, 147-160.
- DiMaggio, Paul; Powell, Walter (1991) “Introduction: The New Institutionalism and Organizational Analysis” in DiMaggio, Paul; Powell, Walter (eds.) *The New Institutionalism and Organizational Analysis*. Chicago: University of Chicago Press.
- Dobrai, Katalin; Farkas, Ferenc (2010) “Growing Professionalism in the Services of Nonprofit Organizations” *Review of International Comparative Management*, 11 (5), 894–904.
- Drori, Gili S.; Meyer, John W.; Hwang, Hokyu (eds.) (2006) *Globalization and organization: World society and organizational change*. New York: Oxford University Press.

- Drucker, Peter; Rangan, Kasturi (2008) "What Is Our Plan?" in Druker, Peter (ed.) *The Five Most Important Questions You Will Ever Ask About Your Organization*. San Francisco, CA: Jossey-Bass. 63–76.
- Ebrahim, Alnoor (2002) "Information struggles: The role of information in the reproduction of NGO-funder relations" *Nonprofit and Voluntary Sector Quarterly*. 31, 85-113.
- Ebrahim, Alnoor (2003a) "Accountability In Practice: Mechanisms for NGOs" *World Development*. 31 (5), 813–29.
- Ebrahim, Alnoor (2003b) "Making Sense of Accountability: Conceptual Perspectives for Northern and Southern Nonprofits" *Nonprofit Management and Leadership*. 14 (2), 191–212.
- Ebrahim, Alnoor (2005) "Accountability Myopia: Losing Sight of Organizational Learning" *Nonprofit and Voluntary Sector Quarterly*. 34(1), 56–87.
- Ebrahim, Alnoor (2009) "Placing the Normative Logics of Accountabilityn'Thick' Perspective" *American Behavioral Scientist*. 52 (6), 885–904.
- Ebrahim, Alnoor (2010) "The Many Faces of Nonprofit Accountability" *Jossey-Bass Handbook of Nonprofit Leadership and Management*. 3 (August), 98–121.
- Ebrahim, Alnoor; Rangan, V. Kasturi (2010) "Putting the Brakes on Impact: A Contingency Framework for Measuring Social Performance" *Academy of Management Annual Meeting Proceedings*. 8 (1), 1–6.
- Ebrahim, Alnoor; Rangan, V. Kasturi (2014) "What Impact? A Framework for Measuring the Scale and Scope of Social Performance" *California Management Review*. 56 (3): 118–41.
- Eckerd, Adam; Moulton, Stephanie (2011) "Heterogeneous Roles and Heterogeneous Practices: Understanding the Adoption and Uses of Nonprofit Performance Evaluations" *American Journal of Evaluation*. 32 (1), 98–117.
- Eco, Humberto (2016) *Os Limites da Interpretação*. São Paulo: Editora Perspetiva. (2.^a Edição).
- Edwards, Michael; Fowler, Alan (eds.) (2002) *The Earthscan Reader on NGO Management*. London: Earthscan.
- Edwards, Michael; Hulme, David (1996) "Too Close for Comfort? The Impact of Official Aid on Nongovernmental Organizations" *World Development*. 24 (6), 961–973.
- Edwards, Michael; Hulme, David (eds.) (1995) *Non-governmental organisations - performance and accountability. Beyond the magic bullet*. London: Earthscan.

- Ellis, Jean; Gregory, Tracy (2008) *Developing monitoring and evaluation in the third sector: Research report*. London: Charities Evaluation Service (CES).
- Enjolras, Bernard (2015) "Measuring the Impact of the Third Sector: From Concepts to Metrics", *TSI Working Paper 05*. Brussels: Third Sector Impact, http://thirdsectorimpact.eu/site/assets/uploads/documentations/tsi-working-paper-no-5-measuring-the-impact-of-the-third-sector-from-concept-to-metrics/TSI_WP5_Measuring-TSI_Concept_to_Metrics1.pdf [24 de Junho 2016].
- Enjolras, Bernard (2016) "Assessing the macro socio- economic impact of the third sector in Europe", *TSI Working Paper 10*. Brussels: Third Sector Impact, http://thirdsectorimpact.eu/site/assets/uploads/documentations/tsi-working-paper-no-10-assessing-macro-socio-economic-impact-third-sector-europe/TSI-Working-Paper-No.10_15p-1x.pdf [19 de Março 2017].
- EQUAL (Gabinete de Gestão) (2003) "Auto-Avaliação dos Projectos" [1ª Fase] *Saber Fazer 3*. Lisboa: Gabinete de Gestão EQUAL.
- EQUAL (Gabinete de Gestão) (2004) "Produtos EQUAL: Validar para Disseminar" *Saber Fazer 4*. Lisboa: Gabinete de Gestão EQUAL.
- EQUAL (Gabinete de Gestão) (2005a) "Auto-avaliação dos Projectos" [2.ª Fase] *Saber Fazer 5*. Lisboa: Gabinete de Gestão EQUAL.
- EQUAL (Gabinete de Gestão) (2005b) *Avaliação independente dos Projectos EQUAL – Orientações*. Lisboa: Gabinete de Gestão EQUAL
- EQUAL (Gabinete de Gestão) (2007a) *Validação de Produtos Inovadores*. Lisboa: Gabinete de Gestão EQUAL.
- EQUAL (Gabinete de Gestão) (2007b) *Disseminar Produtos: Referencial de Apoio*. Lisboa: Gabinete de Gestão EQUAL.
- Estêvão, Madalena (2002) "Balanço de Competências" [1ª Fase], *Saber Fazer 1*, Lisboa: Gabinete de Gestão EQUAL.
- Estêvão, Madalena (2005) "Balanço de Competências" [2ª Fase], *Saber Fazer 8*, Lisboa: Gabinete de Gestão EQUAL.
- Evers, Adalbert; Laville, Jean-Louis (2004) "Defining the Third Sector in Europe" in Evers, Adalbert; Laville, Jean-Louis (eds.) *The Third Sector in Europe*, Cheltenham and Northampton: Edward Elgar: 11–42.
- Feller, Irwin (2002) "Performance measurement redux" *American Journal of Evaluation*. 23, 435–452.
- Ferreira, Sílvia (2008) "A questão social e as alternativas da sociedade civil no contexto das novas formas de governação" *Ciências Sociais Unisinos*. 44, 1, 28-38.

- Ferreira, Sílvia (2009) "Governança Local" in Cattani, António; Gaiger, Luis: Hespanha, Pedro; Laville, Jean Louis (org.) *Dicionário Internacional da Outra Economia*. Coimbra: Almedina.
- Ferreira, Sílvia (2011) "Governance and failure in local strategic multi-sectoral partnerships: the cases of LSP (England) and Rede Social (Portugal)" in Enjolras, Bernard; Bozzini, Emanuela (org.) *Governing Ambiguities: New Forms of Governance and Civil Society*. Baden-Baden: Nomos, 39-56.
- Ferreira, Sílvia (2012) "Observando a indecidibilidade da participação do terceiro setor na governação em rede" *Revista Crítica de Ciências Sociais*. 97, 107-132.
- Ferreira, Sílvia (2013) "Terceiro Sector e Estado-Providência em Portugal" in Silva, Filipe C. (org.) *Os Portugueses e o Estado-Providência*. Lisboa: Imprensa de Ciências Sociais, 161-196.
- Ferreira, Sílvia (2015) "New Paths for Third Sector Institutions in a Welfare State in Crisis: The Case of Portugal" *Nonprofit Policy Forum*. 6, 2, 213-241.
- Ferreira, Virgínia; Ferreira, Sílvia; Coelho, Lina; Lopes, Mónica (2007) *Estudo Diagnóstico e Prospectivo do Papel das ONG na promoção das políticas de igualdade*. Coimbra: CES.
- Ferreira, Virgínia; Monteiro, Rosa; Lopes, Mónica; Neto, Hernâni (2013) *Estudo de Avaliação da Integração da Perspetiva do Género nos Fundos Estruturais, no Período de Programação 2007-2013*. Lisboa: IGFSE.
- Figueirôa, David; Pimenta, Manuel (2005) "A Auto-Avaliação nos Projectos EQUAL" *Cidades – Comunidades e Territórios*. 11, 11-20.
- Fine, Allison H.; Thayer, Colette E.; Coghlan, Anne (2000) Program evaluation practice in the nonprofit sector. *Nonprofit Management and Leadership*. 10(3), 331–339.
- Fitzpatrick, Jody L.; Sanders, James R.; Worthen, Blaine R. (2004) *Program evaluation: Alternative approaches and practical guidelines*. Boston: Pearson Education. (4.ª Edição).
- Fleischer, Dreolin; Christie, Christina (2009) "Evaluation Use. Results from a Survey of US American Evaluation Association Members" *American Journal of Evaluation* 30 (2), 158–75.
- Forbes, Daniel P. (1998) "Measuring the Unmeasurable: Empirical Studies of Nonprofit Organization Effectiveness from 1977 to 1997" *Non-profit and Voluntary Sector Quarterly*. 27, 183–202.
- Fowler, Alan (1997) *Striking a balance: a guide to enhancing the effectiveness of non-governmental organisations in international development*. London: Earthscan.

- Fowler, Alan; Malunga, Chiku W. (2010) *NGO Management: The Earthscan Companion*. London/New York: Earthscan.
- Franco, Raquel C. (coord.) (2015) *Diagnóstico das ONG Em Portugal*. Lisboa: Fundação Calouste Gulbenkian.
http://www.gulbenkian.pt/mediaRep/gulbenkian/files/institucional/actividades/programas_projectos/EEAgrants_CidAtiva/Docs/Diagn__stico_das_ONG_em_Portugal.pdf. [4 de Maio de 2016]
- Franco, Raquel; Sokolowski, S. W.; Hairel, E. M.; Salamon, L. M. (2005) *O setor não lucrativo português numa perspetiva comparada*. Porto: Universidade Católica Portuguesa.
- Freeman, Robert E. (1984) *Strategic management: A stakeholder approach*. Boston: Harper Collins.
- Fry, Ronald E. (1995) "Accountability in Organizational Life: Problem or Opportunity for Nonprofits?" *Nonprofit Management and Leadership*. 6 (2), 181–195.
- Furubo, Jan-Eric, Rist, Ray C., and Sandahl, Rolf (eds.) (2002) *International Atlas of Evaluation*. London: Transaction.
- Garbett, Rob; McCormack, Brendan (2001) "The Experience of Practice Development: An Exploratory Telephone Interview Study" *Journal of Clinical Nursing*. 10 (1), 94–102.
- Ghanem, Maurad; Pedra, Cláudia; Matos, Rosa (2016) "Financiamento do Terceiro Setor em Portugal - Diagnóstico do Projeto SPOTS FASE I". http://www.stone-soup.net/files/DIAGNOSTICO_SPOTS_PORTUGAL_fev_2016.pdf [28 de Junho de 2016]
- Godenhjelm, Sebastian; Lundin, Rolf A.; Sjöblom, Stefan, (2015) "Projectification in the public sector – the case of the European Union" *International Journal of Managing Projects in Business*, 8(2), 324–348.
- Grant, Heather M.; Crutchfield, Leslie R. (2007) "Creating High-Impact Nonprofits" *Stanford Social Innovation Review*. Outono, 32–41.
- Greene, Jennifer C. (1988) "Stakeholder participation and utilization in program evaluation" *Evaluation Review*. 12 (3), 91–116.
- Greene, Jennifer C. (1994) "Qualitative Program Evaluation. Practice and Promise" in Denzin, Norman K.; Lincoln, Yvonna S. (eds.), *Handbook of Qualitative Research* (530-44). London: Sage.
- Greene, Jennifer C. (1999) "The Inequality of Performance Measurements. Evaluation", *Evaluation*. 5(2), 160–172.
- Greene, Jennifer C. (2006) "Evaluation, Democracy, and Social Change" in Shaw, Ian; Greene, Jennifer C.; Melvin, M. Mark. (eds.) *The Sage Handbook of Evaluation : Policies, Programs and Practices*. Thousand Oaks: Sage.

- Greene, Jennifer C.; Caracelli, Valerie J; Graham, Wendy F. (1989) "Toward a Conceptual Framework for Mixed-Method Evaluation Designs" *Educational Evaluation and Policy Analysis*. 11 (3), 255–74.
- Guba, Egon G.; Lincoln, Yvonna S. (2005) "Paradigmatic controversies, contradictions, and emerging confluences" *In* Denzin, Norman; Lincoln, Yvonna (eds.) *SAGE handbook of qualitative research*. Thousand Oaks: Sage.
- Guba, Egon. G.; Lincoln, Yvonna S. (1989) *Fourth generation evaluation*. Newbury Park: Sage.
- Guerra, Isabel (2002) *Fundamentos e processos de uma sociologia de ação: o planeamento em ciências sociais*. Cascais: Principia.
- Hall, Matthew (2014) "Evaluation Logics in the Third Sector" *Voluntas*. 25 (2), 307–36.
- Hall, Michael H., Phillips, Susan D.; Pickering, Donna (2003) *Assessing Performance*. Toronto: Canadian Centre for Philanthropy.
- Harlock, Jenny (2013) "Impact Measurement Practice in the UK Third Setor : A Review of Emerging Evidence", *Third Sector Research Centre Working Paper 106*, Birmingham: TSRC, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-106.pdf> [19 de Setembro de 2014].
- Henriques, José M. (2006) *Global Restructuring and Local Anti-Poverty Action: Learning from European Experimental Programmes*. Tese de doutoramento em Economia. Instituto Superior de Ciências do Trabalho e da Empresa. [https://repositorio.iscte-iul.pt/bitstream/10071/273/1/Tese%20de%20Doutoramento Jos%C3%A9%20Manuel%20Henriques.pdf](https://repositorio.iscte-iul.pt/bitstream/10071/273/1/Tese%20de%20Doutoramento%20Jos%C3%A9%20Manuel%20Henriques.pdf) [12 de novembro de 2011].
- Herman, Robert D. (1990) "Methodological issues in studying the effectiveness of nongovernmental and nonprofit organizations" *Nonprofit and Voluntary Sector Quarterly*, 19(3), 293-306.
- Herman, Robert D., & Associates (2005) *The Jossey-Bass handbook of nonprofit leadership and management*. San Francisco: Jossey-Bass. (2.ª Edição).
- Herman, Robert D.; Renz, David O. (1999) "Theses on Nonprofit Organizational Effectiveness" *Nonprofit and Voluntary Sector Quarterly*. 28 (2), 107–26.
- Herman, Robert D.; Renz, David O. (2004) "Doing things right: Effectiveness in local nonprofit organizations, a panel study" *Public Administration Review*. 64(6), 694-703.
- Hespanha, Pedro (2008) "Políticas Sociais: novas abordagens, novos desafios?" *Revista Crítica de Ciências Socais*, 39, 1, 5-15.
- Hofer, Richard (2000) "Accountability in action? Program evaluation in nonprofit human services agencies" *Nonprofit Management and Leadership*. 11(2), 167-177.

- Hoefler, Richard (2009) "Program Evaluation for Accountability : What ' S Different Now ?" *ISTR Conference Working Papers Volume VI*, https://c.ymcdn.com/sites/istr.siteym.com/resource/resmgr/working_papers_barcelona/Hoefler.pdf [9 de Setembro de 2010].
- Højlund, Steven (2014) "Evaluation Use in the Organizational Context - Changing Focus to Improve Theory" *Evaluation*. 20, 26–43.
- Hood, Christopher (1991) "A public management for all seasons?" *Public Administration*. 69(1), 3–19.
- Hoole, Emily; Patterson, Tracy E. (2008) "Voices from the field: Evaluation as part of a learning culture", *Nonprofits and Evaluation: New Directions for Evaluation*. 93–113.
- Horváth, Péter (2011) *Controlling*. München: Vahlen. (12.ª Edição).
- Huber, Manfred (2007) "Long-term care for older people: The future of Social Services of General Interest in the European Union" *Discussion paper for the Peer Review on Social Protection and Social Inclusion*. Bélgica.
- Hudson, Michael (2003) *Managing at the leading edge*. San Francisco: Jossey-Bass.
- Hvenmark, Johan (2013) Business as Usual? *Administrative Theory & Praxis*. 35(2), 223–247.
- Hwang, Hokyung; Powell, Walter W. (2009) "Charity: The Influences of Professionalism in the Nonprofit Sector" *Administrative Science Quarterly* 54: 268–98.
- Innovation Network (2010) *State of evaluation 2010: Evaluation practice and capacity in the nonprofit sector*. Innovation Network. http://www.innonet.org/client_docs/innonet-state-of-evaluation-2010.pdf. [9 de Abril de 2011].
- Innovation Network (2012) *State of evaluation 2012: Evaluation practice and capacity in the nonprofit sector*. Innovation Network. <https://www.innonet.org/media/innonet-state-of-evaluation-2012.pdf> [19 de Janeiro de 2013].
- Innovation Network (2016) *State of evaluation 2016: Evaluation practice and capacity in the nonprofit sector*. Innovation Network, https://www.innonet.org/media/2016-State_of_Evaluation.pdf. [6 de Janeiro de 2017].
- Instituto de Gestão Financeira da Segurança Social (IGFSS) (2008) CONTA DA SEGURANÇA SOCIAL 2008. Lisboa: IGFF. www.seg-social.pt/documents/10152/48574/conta_seguranca_social_2008/ba2a1883-5097-4e22-adca-53a7edc2c6e8 [18 de Outubro de 2016]
- Instituto de Gestão Financeira da Segurança Social (IGFSS) (2014) CONTA DA SEGURANÇA SOCIAL 2014. Lisboa: IGFSS. http://www.seg-social.pt/documents/10152/48574/conta_seguranca_social_2014/ba2a1883-5097-4e22-adca-53a7edc2c6e8

social.pt/documents/10152/13433747/Conta+da+Seguran%C3%A7a+Social+de+2014+-+Parte+I [18 de Outubro de 2016]

Instituto de Segurança Social (ISS, IP) (2007) Manual/Linhas orientadoras do Apoio Técnico às IPSS com Acordo de Cooperação. Lisboa: ISS, IP.

Instituto Nacional de Estatística (INE) (2011) Conta Satélite das Instituições sem fim lucrativo. Lisboa: INE, https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=122650440&att_display=n&att_download=y [27 de Fevereiro de 2014].

Instituto Nacional de Estatística (INE) e Cooperativa António Sérgio para a Economia Social (CASES) (2013) *Conta Satélite da Economia Social - 2010*. Lisboa: INE; CASES, [http://www.cases.pt/wp-content/uploads/2016/12/Destaque Conta Satelite da Economia Social.pdf](http://www.cases.pt/wp-content/uploads/2016/12/Destaque%20Conta%20Satelite%20da%20Economia%20Social.pdf) [30 de Março de 2015].

Instituto Nacional de Estatística (INE) e Cooperativa António Sérgio para a Economia Social (CASES) (2013) *Conta Satélite da Economia Social – 2013*. Lisboa: INE; CASES, [http://www.cases.pt/wp-content/uploads/2016/12/Destaque Conta Satelite da Economia Social.pdf](http://www.cases.pt/wp-content/uploads/2016/12/Destaque%20Conta%20Satelite%20da%20Economia%20Social.pdf) [4 de Fevereiro de 2017].

Jacobs, Alex; Barnett, Chris; Ponsford, Richard (2010) “Three approaches to monitoring: Feedback systems, participatory monitoring and evaluation and logical frameworks”, *IDS Bulletin*. 41, 36–44.

Jailobaeva, Kanykey (2014) “A New Look: Professionalization of NGOs in Kyrgyzstan” *Central Asian Survey*. 33 (3), 360–74.

Jalocha, Beata (2012) “Projectification of the European Union and Its Implications for Public Labour Market Organisations in Poland”, *Journal of Project, Program & Portfolio Management*. 3 (2), 1–16.

Jessop, Bob (2003) “Governance and Metagovernance: On reflexivity, Requisite Variety, and Requisite Irony” in Bang, Henrik P. (org.), *Governance as Social and Political Communication*. Manchester: Manchester University Press, 142-172.

John, Rob (2006) “Venture Philanthropy: The Evolution of High Engagement Philanthropy in Europe” *The Evolution of High Engagement Philanthropy in Europe*. Junho, 44.

Johnsen, Åge (2013) “Performance management and evaluation in Norwegian local government: Complementary or competing tools of management?” *New Directions for Evaluation*. 137, 93–101.

Johnson, Kelli; Toal, Stacie A.; King, Jean A.; Volkov, Boris (2009) Research on Evaluation Use *American Journal of Evaluation*. 30(3), 377–410.

- Kaplan, Robert S. (2001) "Strategic performance measurement and management in nonprofit organizations" *Nonprofit Management & Leadership*. 11, 353-370.
- Kaplan, Robert S; Norton, David P. (1992) "The balanced scorecard: Measures that drive performance" *Harvard Business Review*. 70(1), 71-79.
- Kaplan, Robert S; Norton, David P. (1996) *The Balanced Scorecard Translating Strategy In Action*. Boston: Harvard Business School Press.
- Kaplan, Robert S; Norton, David P. (2004) *Strategy maps: Converting intangible assets into tangible outcomes*. Boston: Harvard Business School Press.
- Kearns, Kevin P. (1996) *Managing for Accountability: Preserving the Public Trust in Nonprofit Organizations*. San Francisco: Jossey-Bass.
- Kendall, Jeremy; Knapp, Martin (2000) "Measuring the Performance of Voluntary Organizations" *Public Management Review* 2 (1), 105–32.
- King, Jean; Stevahn, Laurie (2013) *Interactive Evaluation Practice: Mastering the Interpersonal Dynamics of Program Evaluation*. Los Angeles: Sage.
- Kloos, Karina; Papi, Daniela (2014) "Lost in Translation" *Stanford Social Innovation Review*, Verão. <https://communityengagement.uncg.edu/wp-content/uploads/2014/08/Lost-In-Translation.pdf> [28 de Outubro de 2015].
- Korff, Valeska P; Oberg, Achim; Powell, Walter W. (2015) "Interstitial Organizations as Conversational Bridges" *Bulletin of the Association for Information Science and Technology*. 41 (2), 34–38.
- Kosecoff, Jacqueline. E; Fink, Arlene (1982) *Evaluation Basics. A Practitioner's Manual*. Thousand Oaks: Sage.
- Kováč, Imre; Eva, Kučerová (2006) "The project class in Central Europe: the Czech and Hungarian cases" *Sociologia Ruralis*. 46 (1), 3-21.
- Kuura, Arvi (2011) "Policies for projectification: support, avoid or let it be? *Discussions on Estonian Economic Policy*. <http://ssrn.com/abstract=1884204> [5 de Junho de 2014].
- Lahey, Robert, Nielsen, Steffen B. (2013) Rethinking the relationship among monitoring, evaluation and results-based management: Observations from Canada" *New Directions for Evaluation*. 137, 45–56.
- Lecy, Jesse D.; Schmitz, Hans P.; Swedlund, Haley (2012) "Non-Governmental and Not-for-Profit Organizational Effectiveness: A Modern Synthesis" *Voluntas*. 23, 434–57.
- Lewis, David (2001) *The management of non-governmental development organizations: an introduction*. London/New York: Routledge.

- Liket, Kellie C.; Rey-Garcia, Marta; Maas, Karen E. H. (2014) "Why Aren't Evaluations Working and What to Do About It: A Framework for Negotiating Meaningful Evaluation in Nonprofits" *American Journal of Evaluation*. 35: 171–88.
- Lincoln, Yvonna S; Guba, Egon G. (2013) The Roots of Fourth Generation Evaluation: Theoretical and Methodological Origins in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: Sage, 218-228. (2.ª Edição).
- Lindenberg, Marc; Bryant, Coralie (eds.) (2001) *Going Global: Transforming Relief and Development NGOs*. Bloomfield, CT: Kumarian Press.
- Lopes, Alexandra; Parente, Cristina; Marcos, Vanessa (2014) "Entre o Estado e o Terceiro Setor: modos de regulação. O Terceiro Setor português em foco" in Parente, Cristina (ed.) *Empreendedorismo Social em Portugal*. Porto: Universidade do Porto, 75-96.
- Lucas, Susana; Sardinha, Boguslawa; Dias Otília; Pires, António (2012) "Sistemas de certificação da qualidade nas instituições de respostas sociais" *TMQ – Techniques, Methodologies and Quality Review*. [Vol.?, n.º?]. Instituto Politécnico de Setúbal. https://comum.rcaap.pt/bitstream/10400.26/5314/1/Artigo_Certifica%C3%A7%C3%A3o%20de%20qualidade%20nas%20respostas%20sociais_18-05-2013.pdf [23 de Junho de 2015].
- Lumley, Tris; Rickey, Benedict; Pike, Matthew (2011) *Inspiring impact: Working together for a bigger impact in the UK social sector*. London: NPC.
- Lundstrom, Tommy (2001) "Child protection, voluntary organizations, and the public sector in Sweden" *Voluntas*. 12, 355-371.
- Lyon, Fergus; Arvidson, Malin (2011) "Social Impact Measurement as an Entrepreneurial Process" *Third Sector Research Centre Briefing Paper 66*, Birmingham: TSRC, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/briefing-paper-66.pdf> [10 de Janeiro de 2012]
- MacIndoe, Heather; Barman, Emily (2012) "How Organizational Stakeholders Shape Performance Measurement in Nonprofits: Exploring a Multidimensional Measure" *Nonprofit and Voluntary Sector Quarterly*. 42 (2), 716–38.
- Marôco, João (2014) *Análise Estatística com o SPSS Statistics*. Pêro Pinheiro: ReportNumber.
- Marshall, Tony F. (1996) "Can we define the voluntary sector?" In Billis, David; Harris, Margaret (eds.) *Voluntary agencies: Challenges of organization and management*, London: Macmillan.
- Martín, Ignacio; Mota, Jorge; Santos, Natacha; Silva, Sílvia (2014) *Sustentabilidade das Instituições Particulares de Solidariedade Social em Portugal*, [S.l., s.n].

- Martz, Wes (2013) "Evaluating Organizational Performance: Rational, Natural, and Open System Models" *American Journal of Evaluation*. 34 (3), 385–401.
- Mass, Flipe; Rodrigues, Ricardo (2010) "Achieving Excellence in Social Service Provision A . The Policy Debate at European Level", *Discussion Paper for the Peer Review on Social Protection and Social Inclusion*. Bélgica.
- Mayne, John (2010) "Building an Evaluative Culture: The Key to Effective Evaluation and Results Management" *Canadian Journal of Program Evaluation*. 24 (2), 1–30.
- Mayne, John; Rist, Ray C. (2006) "Studies are not enough: The necessary transformation of evaluation" *The Canadian Journal of Program Evaluation*, 21, 93–120.
- McDavid, James C., Huse, Irene; Hawthorn, Laura. (2013) *Program evaluation and performance measurement: An introduction to practice*. Thousand Oaks: Sage. (2.^a Edição).
- McDonald, Catherine (1997) "Government, funded nonprofits, and accountability" *Nonprofit Management and Leadership*. 8(1), 51–64.
- Melloni, Erica (2012) *Come le organizzazioni Nonprofit valutano l'impatto delle proprie attività* (Relatório de pesquisa). Milão: Fondazione Sodalitas/Instituto per la Ricerca Social, http://www.sodalitas.it/public/allegati/INDAGINE_RISULTATI_201531512283440.pdf [4 de Junho 2014].
- Mertens, Donna M. (2009) *A Transformative Research and Evaluation*. New York: Guilford Press.
- Mertens, Donna M. (2010) *Research and Evaluation in Education and Psychology: Integrating Diversity With Quantitative, Qualitative, and Mixed Methods*. Thousand Oaks: Sage.
- Mertens, Donna M; Wilson, Amy (2012) *Program Evaluation Theory and Practice: A Comprehensive Guide*. New York: Guilford Press.
- Mesquita, João (2015) *Cultura de Avaliação no Setor da Cooperação para o Desenvolvimento em Portugal*. Dissertação de Mestrado em Desenvolvimento e Cooperação Internacional. Universidade de Lisboa. <https://www.iseg.ulisboa.pt/aquila/getFile.do?fileId=639486&method=getFile> [30 de Maio de 2016]
- Meyer, John; Rowan, Brian (1977) "Institutional organizations: Formal structure as myth and ceremony" *American Journal of Sociology*. 83, 340-363.

- Millar, Ross; Kelly, Hall, (2013) "Social return on investment (SROI) and performance measurement: The opportunities and barriers for social enterprises in health and social care" *Public Management Review*. 15(6), 923–941.
- Mitchell, George E. (2014) "Why Will We Ever Learn? Measurement and Evaluation in International Development NGOs" *Public Performance & Management Review*. 37 (4), 605–631.
- Mitchell, George E.; Berlan, David (2016) "Evaluation and Evaluative Rigor in the Nonprofit Sector" *Nonprofit Management and Leadership*. 27 (2), 237–250.
- Monteiro, Rosa (2011) *Feminismo de Estado Em Portugal: Mecanismos, Estratégias, Políticas e Metamorfoses*. Tese de Doutoramento em Sociologia. Faculdade de Economia.
- Moody, Michael (2008) "'Building a Culture': The Construction and Evolution of Venture Philanthropy as a New Organizational Field" *Nonprofit and Voluntary Sector Quarterly*. 37 (2), 324–52.
- Moseley, James L.; Dessinger, Joan C. (eds.) (2010) *Handbook of Improving Performance in the Workplace: Measurement and Evaluation*. San Francisco: Wiley.
- Moynihan, Donald (2008) *The dynamics of performance management: Constructing information and reform*. Washington, DC: Georgetown University Press.
- Murray, Vic (2005) "Evaluating the effectiveness of nonprofit organizations" in Herman, Robert (ed.), *The Jossey-Bass handbook of nonprofit leadership and management*. San Francisco: Jossey-Bass, 345-370. (2.ª Edição).
- Murray, Vic; Bradshaw, Pat; Wolpin, Jacob (1992) "Power in and Around Nonprofit Boards: A Neglected Dimension of Governance" *Nonprofit Management and Leadership*. 3, 165–182.
- Murray, Vic; Tassie, Bill (1994) Evaluating the effectiveness of nonprofit organizations. in Herman, Robert & Associates, *The Jossey-Bass handbook of nonprofit leadership and management*. 303–324. San Francisco: Jossey-Bass.
- Musselwhite, Kimberly; Cuff, Laura; McGregor, Lisa; King, Kathryn M. (2007) "The Telephone Interview Is an Effective Method of Data Collection in Clinical Nursing Research: A Discussion Paper" *International Journal of Nursing Studies*. 44 (6), 1064–1070.
- Najam, Adil (1996) "NGO accountability: A conceptual framework" *Development Policy Review*. 14, 339-353.
- Newcomer, Kathryn E. (2004) How might we strengthen evaluation capacity to manage evaluation contracts? *American Journal of Evaluation*. 25(2), 209–218.

- Nicolau, Isabel; Simaens, Ana (2008) Strategic Management in Social Economy - an overview of Social Solidarity Cooperatives in Portugal. *CIRIEC Working Paper 2008/06* <http://www.ciriec.ulg.ac.be/wp-content/uploads/2015/11/WP08-06.pdf> [6 de Novembro de 2010].
- Nielsen, Steffen B.; Hunter David (2013) "Challenges to and Forms of Complementarity Between Performance Management and Evaluation" *New Directions for Evaluation*. 137, 115–23.
- Nielsen, Steffen B; Ejler, Nicolaj (2008) "Improving Performance?: Exploring the Complementarities between Evaluation and Performance Management" *Evaluation*. 14 (2), 171–92.
- Novick, Gina (2008) "Is There a Bias against Telephone Interviews in Qualitative Research?" *Research in Nursing and Health*. 31 (4), 391-398.
- Oelberger, Carrie; Oberg, Achim; Kloos, Karina; Korff, Valeska; Powell, Woody (2012) "Co-exist, Colonize, or Combine? Accounting for patterns of discourse on nonprofit evaluation" *In EGOS (European Group for Organizational Studies) 28th EGOS Colloquium*. Helsínquia, Finlândia 5-7 de julho de 2012. <https://pt.slideshare.net/eFogg/egos-presentation-for-pacs-website-valeska> [28 de Setembro de 2014].
- Ógáin, Eibhlín N.; Lumley, Tris; Pritchard, David (2012) *Making an Impact. Impact Measurement among Charities and Social Enterprises in the UK*. London: Charities Evaluation Services.
- Osborne, David; Gaebler, Ted (1993) "Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector" *Resenhas*. 33(6), 97–99.
- Osborne, Stephen P. (1992) "The quality dimension: Evaluating quality of service and quality of life in human services" *British Journal of Social Work*. 22(4), 437-453.
- Owen, John M.; Lambert, Faye C. (1998) "Evaluation and the Information Needs of Organizational Leaders" *American Journal of Evaluation*, 19 (3), 355–365.
- Packendorff, Johann; Lindgren, Monica (2014) "Projectification and its consequences: Narrow and broad conceptualisations" *South African Journal of Economic and Management Sciences*. 17 (1), 7-2.
- Parente, Cristina (2014) (ed.) *Empreendedorismo Social em Portugal*. Porto: Universidade do Porto.
- Parente, Cristina; Quintão, Carlota (2014) "Uma abordagem eclética ao empreendedorismo social" *in* Parente, Cristina (ed.) *Empreendedorismo Social em Portugal*. Porto: Universidade do Porto.

- Paton, Rob (2003) *Managing and measuring social enterprises*. Thousand Oaks: Sage.
- Patton, Michael (2013) "The Roots of Utilization-Focused Evaluation" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, 293-303. (2.^a Edição).
- Patton, Michael Q. (1978) *Utilization-Focused Evaluation*. Thousand Oaks: Sage.
- Patton, Michael Q. (1997) *Utilization-focused evaluation*. Thousand Oaks: Sage. (3.^a Edição).
- Patton, Michael Q. (2008) *Utilization-focused evaluation*. Thousand Oaks: Sage. (4.^a Edição).
- Patton, Michael Q. (2012) *Essentials of Utilization-Focused Evaluation*. Thousand Oaks: Sage.
- Perrin, Burt (2006) *Moving from Outputs to Outcomes: Practical Advice from Governments around the World*. Arlington, VA: IBM Center for the Business of Government.
- Pfeffer, Jeffrey; Salancik, Gerald R. (2003) *The external control of organizations: A resource dependence perspective*. New York: Harper & Row.
- Pierre, Jon (2000) "Understanding Governance", in Pierre, Jon (org.) *Debating Governance: Authority, Steering and Democracy*, Oxford: Oxford University Press.
- Piotrowski, Suzanne J.; Rosenbloom, David H. (2002) "Nonmission-based values in results-oriented public management: The case of freedom of information" *Public Administration Review*. 62, 643–657.
- Poister, Theodore (2010) "Performance measurement: Monitoring program outcomes" in Wholey, Joseph; Hatry, Harry; Newcomer, Kathryn (eds.), *Handbook of practical program evaluation*. San Francisco: Jossey-Bass, 100-124. (3.^a Edição).
- Poister, Theodore; Aristigueta, M. P.; Hall; J. L. (2014) *Managing and Measuring Performance in Public and Nonprofit Organizations: An Integrated Approach*. Hoboken: Wiley. (2.^a Edição).
- Pollitt, Christopher; Summa, Hilikka (1997) "Performance Auditing" in Chelimsky, Eleanor; Shadish, William (eds.) *Evaluation for the 21st Century*. Thousand Oaks: Sage, 99-122.
- Posavac, Emil J.; Carey, Raymond (2003) *Program Evaluation: Methods and Cases*. Upper Saddle River, NJ: Prentice Ha.
- Powell, Ellen; Boyd, Heather H. (2008) "Evaluation Capacity Building in Complex Organizations" *New Directions for Evaluation*. 2008 (120), 55–69.
- Preskill, Hallie T.; Torres, Rosalie T. (1999) *Evaluative inquiry for learning in organizations*. Thousand Oaks: Sage.

- Preskill, Hallie T.; Torres, Rosalie T. (2000) "The Learning Dimension of Evaluation Use" *New Directions for Evaluation*, 88, 25–37.
- Preskill, Hallie; Boyle, Shanelle (2011) "Insights into Evaluation Capacity Building: Motivations, Strategies, Outcomes, and Lessons Learned" *The Canadian Journal of Program Evaluation*. 23 (3), 147–74.
- Programa Operacional Potencial Humano (POPH) (2014) Programa Operacional Temático Potencial Humano POPH 2007-2013. http://www.poph.qren.pt/upload/docs/apresenta%C3%A7%C3%A3o/Programa/PO_Potencial_Humano_Fevereiro2014.pdf [17 de Março de 2014].
- Programa Operacional Potencial Humano (POPH) (2015) Relatório de Execução 2014. <http://www.poph.qren.pt/upload/docs/%C3%81rea%20Reservada/2015/Relatorio Execu%C3%A7%C3%A3o%20 F POPH 2014.pdf> [22 de Janeiro de 2016].
- Puyvelde, Stijn; Caers, Ralf; Du Bois, Cind; Jegers, Marc (2012) "The Governance of Nonprofit Organizations: Integrating Agency Theory with Stakeholder and Stewardship Theories" *Nonprofit and Voluntary Sector Quarterly*. 41 (3), 431–51.
- Quinn, Robert E.; Rohrbaugh, John (1983) "A spatial model of effectiveness criteria: Towards a competing values approach to organizational analysis" *Management Science*, 29, 363-377.
- Quintão, Carlota (2011) "O Terceiro Setor e a sua renovação em Portugal. Uma abordagem preliminar" *IS Working Paper 2*. Porto: Universidade do Porto, Instituto de Sociologia, http://www.isociologia.pt/App_Files/Documents/is_wp_ns_002_110411025242.pdf. [6 de julho de 2013].
- Radin, Beryl (2006) *Challenging the performance movement: Accountability, complexity, and democratic values*. Washington, DC: Georgetown University Press.
- Renz, David (2016) "The Future of Nonprofit Leadership and Management" in Renz, David; Herman, Robert (eds.) *The Jossey-Bass Handbook of Nonprofit Leadership and Management*, New Jersey: Jossey-Bass. 734-745. (4.ª Edição).
- Renz, David; Herman, Robert (2016) "Understanding Nonprofit Effectiveness" in Renz, David; Herman, Robert (eds.) *The Jossey-Bass Handbook of Nonprofit Leadership and Management*. New Jersey: Jossey-Bass. 274-291. (4.ª Edição).
- Rogers, Patricia; Williams, Bob (2006) "Evaluation for Practice Improvement and Organizational Learning" in Shaw, Ian; Greene, Jennifer C.; Melvin, M. Mark. (eds.) (2006) *The Sage Handbook of Evaluation : Policies, Programs and Practices*. Thousand Oaks: Sage. 76-97.

- Rossi, Peter (2013) "My Views of Evaluation and Their Origins" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: Sage, 106-112. (2.ª Edição).
- Ruela, Andreia G. (2015) *Accountability em Instituições Particulares de Solidariedade Social: um estudo exploratório*. Dissertação de Mestrado em Intervenção Social, Inovação e Empreendedorismo. Universidade de Coimbra.
- Ruela, Andreia G.; Albuquerque, Cristina P. (2016) *Accountability no Terceiro Setor em Portugal: Perspetivas, Desafios e Oportunidades*. Coimbra: Imprensa da Universidade de Coimbra.
- Salamon, Lester M. (2005) "The Changing Context of American Nonprofit Management" in Herman, Robert, & Associates (eds.) *The Jossey-Bass Handbook of Nonprofit Leadership and Management*, San Francisco: Jossey-Bass, 81-100. (2.ª Edição).
- Salamon, Lester M.; Geller, Stephanie L.; Mengel, Kasey L. (2010) "Nonprofits, Innovation, and Performance Measurement: Separating Fact from Fiction" *Johns Hopkins Listening Post Project*. 17, 1–23.
- Salamon, Lester, Anheier, H.; List, R.; Toepler, S.; Sokolowski, W., & Associates (1999) *Global civil society: dimension of the nonprofit sector*. Baltimore: Johns Hopkins University, <http://ccss.jhu.edu/wp-content/uploads/downloads/2011/08/Global-Civil-Society-I.pdf> [19 de Setembro de 2011].
- Salamon, Lester; Sokolowski, S. W.; Haddock, M.; Tice, H. S. (2012) *Portugal's nonprofit sector in comparative context*. Baltimore: Johns Hopkins Center for Civil Society Studies, http://ccss.jhu.edu/wp-content/uploads/downloads/2012/04/Portugal_Comparative-Report_FINAL_4.2012.pdf [19 de Dezembro de 2012].
- Santos, Elizabeth; Veiga, Fátima, Cruz, Paula; Lopes, Mónica; Ferreira, Sílvia (2015) *O impacto social e institucional da crise económica e financeira nas Organizações do Terceiro Sector*, Lisboa: EAPN.
- Schein, Edgar H. (1996) "Three Cultures of Management: The Key to Organisational Learning" *Sloan Management Review*. 38 (1), 9–20.
- Schmidt, Vivien A. (2013) "Democracy and Legitimacy in the European Union Revisited: Input, Output and 'Throughput'" *Political Studies*. 61, 1, 2-22.
- Scriven, Michael (2013) "Conceptual Revolutions in Evaluation: Past, Present and Future" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks, 167-179. (2.ª Edição).

- Secção Especializada de Avaliação do Sistema de Controlo Interno da Administração Financeira do Estado (SEAV) (s/d) Glossário de Conceitos de Avaliação e Gestão. Lisboa: Ministério da Finanças, <http://www.ccas.min-financas.pt/documentacao/Glossario-seav.pdf> [18 de Março de 2013].
- Sempik, Joe; Becker, Saul; Bryman, Alan (2007) “The Quality of Research Evidence in Social Policy: Consensus and Dissension among Researchers” *Evidence and Policy*. 3, 407–423.
- Shadish, William R.; Cdeesook, Thomas; Leviton, Laura C. (1991) *Foundations of program evaluation: Theories of practice*. Newbury Park: Sage.
- Sharman, Jason C. (2007) “Rationalist and Constructivist Perspectives on Reputation” *Political Studies*. 55: 20–37.
- Shaw, Ian; Greene, Jennifer C.; Melvin, M. Mark. (eds.) (2006) *The Sage Handbook of Evaluation : Policies, Programs and Practices*. Thousand Oaks: Sage.
- Shiefer, Ulrich; Bal-Döbel, Lucínia; Batista, António; Döbel, Reinald; Nogueira, João; Teixeira, Paulo (2006) *MAPA - Manual de Planeamento e Avaliação de Projetos*. Estoril: Principia.
- Shulha, Lyn; M., J. Bradley, Cousins (1986) “Evaluation Use: Theory, Research, and Practice Since 1986” *Evaluation Practice*. 18(3), 195-208.
- Simister, Nigel; Smith, Rachel (2010) “Monitoring and Evaluating Capacity Building : Is It Really That Difficult ?” *Praxis Paper* 357. Oxford: INTRAC. <http://www.intrac.org/data/files/resources/677/Praxis-Paper-23-Monitoring-and-Evaluating-Capacity-Building-is-it-really-that-difficult.pdf> [15 de Janeiro de 2012].
- Sowa, Jessica E.; Selden, Sally C.; Sandfort, Jodi R. (2004) “No Longer Unmeasurable? A Multidimensional Integrated Model of Nonprofit Organizational Effectiveness” *Nonprofit and Voluntary Sector Quarterly*. 33 (4), 711–28.
- Speckbacher, Gerhard (2003) “The economics of performance management in nonprofit organizations” *Nonprofit Management and Leadership*. 13(3), 267-281.
- Stake, Robert E. (2001) “Representing quality in evaluation” in Benson, Alexis P.; Hinn, D. Michelle; Lloyd, Claire (eds.) *Vision of Quality: How Evaluators Define, Understand and Represent Program Quality* (Advances in Program Evaluation, Volume 7), Emerald Group, 3 – 11.
- Stern, Elliot (2006) “Contextual challenges for evaluation practice” In Shaw, Ian; Greene, Jennifer C.; Melvin, M. Mark (eds.) *The Sage Handbook of Evaluation : Policies, Programs and Practices*. Thousand Oaks: Sage.

- Stewart, Amanda J. (2014) "Understanding Nonprofit Professionalization: Common Concepts and New Directions for Research" *Paper presented at the Academy of Management*. Philadelphia, PA. 1–27, <http://socialimpactstrategy.org/wp-content/uploads/2016/03/Stewart.pdf> [15 de Maio de 2015].
- Stockmann, Reinhard (2011) *A practitioner handbook on evaluation*. Cheltenham: Edward Elgar.
- Stufflebeam, Daniel L. (2002) *Institutionalizing evaluation checklist*. Evaluation Center, Western Michigan University, http://www.wmich.edu/evalctr/archive_checklists/institutionalizingeval.pdf. [12 de junho de 2013].
- Stufflebeam, Daniel L.; Coryn, Chris (2014) *Evaluation Theory, Models, and Applications. Research Methods for the Social Sciences*. San Francisco: Jossey-Bass & Pfeiffer Imprints, Wiley. (2.^a Edição).
- Sturges, Judith E., Hanrahan, Kathleen J. (2004) "Comparing telephone and face-to-face qualitative interviewing: a research note" *Qualitative Research*, 4 (1), 107–118.
- Tassie, Bill; Murray, Vic; Cutt, James (1998) "Evaluating Social Service Agencies: Fuzzy Pictures of Organizational Effectiveness" *Voluntas*. 9 (1), 59–79.
- Torres, Rosalie T.; Preskill, Hallie T. (2001) "Evaluation and Organizational Learning: Past, Present, and Future" *American Journal of Evaluation*. 22 (3), 387–95.
- Tyler, Ralph W. (1949) *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press.
- Vale, Ana (2010) "Resultados do Programa EQUAL em Portugal 2000 – 2010" *Revista Sociedade e Trabalho*. 41, 287-312.
- Vedung, Evert (2004) "Evaluation research and fundamental research" in Stockmann, Reinhard (ed.), *Evaluationsforschung*, Opladen: Leske & Budrich, 111–134. (2.^a Edição)
- Vedung, Evert (2010) "Four Waves of Evaluation Diffusion" *Evaluation*. 16 (3), 263-277.
- Verbruggen, Sandra; Christiaens, Johan; Milis, Koen (2011) "Can Resource Dependence and Coercive Isomorphism Explain Nonprofit Organizations' Compliance With Reporting Standards?" *Nonprofit and Voluntary Sector Quarterly*. 40: 5–32.
- Vernis, Alfred; Iglesias, Maria; Sanz, Beatriz; Saz-Carranza, Àngel (2006) *Nonprofit Organizations: Challenges and Collaboration*. London; Palgrave MacMillan.
- Volkman, Robin (1999) "Outcomes measurement: The new accounting standard for service organizations" *Fundraising Management*. 30, 26–27.

- Volkov, Boris B.; King, Jean A. (2007) "A Checklist for Building Organizational Evaluation Capacity" *The Evaluation Center Western Michigan University*, 1–5, http://www.wmich.edu/evalctr/archive_checklists/ecb.pdf. [26 de Julho de 2012]
- W.K. Kellogg Foundation (2010) *W.K. Kellogg Foundation Evaluation Guide*. Battle Creek, MI: W.K. Kellogg Foundation.
- Weiss, Carol (1981) "Doing research or doing policy" *Evaluation and Program Planning*. 4, 297-402.
- Weiss, Carol (2013) "Rooting for Evaluation: Digging Into Beliefs" in Alkin, Marvin (ed.) *Evaluation Roots – A Wider Perspective of Theorists' Views and Influences*. Thousand Oaks: Sage, 130-143. (2.ª Edição).
- Westall, Andrea (2009) "Value and the Third Sector Working Paper on Ideas for Future Research" *Third Sector Research Centre Working Paper 25*, Birmingham: TSRC, <http://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-25.pdf> [10 de Janeiro de 2012].
- Wholey, Joseph S.; Hatry, Harry P.; Newcomer, Kathryn (eds.) (2010) *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass, 100-124. (3.ª Edição).
- Worth, Michael (2014) *Nonprofit Management: Principles and Practice*. Thousand Oaks: Sage. (3.ª Edição).