

MÉRITO, DESEMPENHO E RESULTADOS: OS DESAFIOS DA MEDIÇÃO

MERIT, PERFORMANCE AND RESULTS: THE CHALLENGES OF MEASUREMENT*

RICARDO RAMOS PINTO**
PEDRO MIGUEL ALVES RIBEIRO CORREIA***
UNIVERSIDADE DE LISBOA, PORTUGAL

Resumo: Este artigo, dedicado à medição do mérito e aos desafios suscitados por essa mesma medição, encontra-se dividido em três secções. Numa primeira secção é feita uma apresentação dos vários tipos de indicadores disponíveis para levar a cabo essa tarefa. A secção seguinte é dedicada aos desafios da medição do mérito no recrutamento e seleção, sendo apresentado um estudo de caso referente à Comissão de Recrutamento e Seleção para a Administração Pública (CReSAP), de Portugal. Finalmente, a terceira secção é dedicada aos desafios da medição do mérito no desempenho, fazendo incidir o foco no Sistema Integrado de Gestão e Avaliação do Desempenho da Administração Pública (SIADAP), sendo apresentados dois estudos de caso referentes aos desafios da medição do mérito no desempenho nos setores do ensino superior e da justiça.

Palavras-chave: Desempenho. Avaliação. Mérito. Administração Pública. Políticas Públicas.

Abstract: This article, dedicated to the measurement of merit and the challenges raised by this same measurement, is divided into three sections. In a first section, the various types of indicators available to carry out this task are presented. The next section is devoted to the challenges of measuring merit in recruitment and selection, and a case study is presented for the Portuguese Committee for Public Administration Recruitment and Selection (CRESAP). Finally, the third section is devoted to the challenges of measuring merit in performance, focusing on the Integrated System of Public Administration Performance Management and Evaluation (SIADAP), presenting two case studies on the challenges of measuring merit in performance in the higher education and justice sectors.

Keywords: Performance. Evaluation. Merit. Public Administration. Public Policy.

* Artigo recebido em 03/09/2018 e aprovado para publicação pelo Conselho Editorial em 30/11/2018.

** Doutoramento em Engenharia Química pelo Instituto Superior Técnico da Universidade Técnica de Lisboa. Curriculum CiênciaVitae: <https://www.cienciavitae.pt/portal/141E-F88D-318D>. E-mail: rpp@iscsp.ulisboa.pt.

*** Doutoramento em Ciências Sociais (Especialidade em Administração Pública) no Instituto Superior de Ciências Sociais e Políticas da Universidade Técnica de Lisboa. Curriculum CiênciaVitae: <https://www.cienciavitae.pt/4914-5E4E-AF20>. E-mail: pcorreia@iscsp.ulisboa.pt.

1. Sobre Escalas de Medida

Para que seja possível medir a ocorrência de uma determinada característica é necessário começar por definir a escala de medida a utilizar, a qual deverá permitir registar todas as variações que a característica em causa possa apresentar. A escala de medida terá que respeitar a natureza dos dados e ser capaz de acomodar todas as particularidades da característica em causa, o que, devido à grande diversidade e complexidade da realidade, resulta numa quantidade quase infinita de escalas. Contudo, é possível agregar todas as escalas de medida em duas grandes famílias, as qualitativas e as quantitativas, sendo que as primeiras se subdividem em nominais e ordinais e as segundas em intervalares e de razão (Ramos Pinto, 2012). A definição da escala de medida é determinante para a qualidade da mediação e da natureza dos indicadores que daí resultam, em termos da exatidão e do grau de detalhe com que a informação é registada, o que condiciona a forma de analisar os dados e o detalhe e abrangência das conclusões.

Imagine-se que no âmbito de um estudo sobre os colaboradores de uma determinada empresa se recolhem diversas informações sobre os mesmos. Se uma das informações pretendidas for o departamento a que determinado colaborador pertence, por mais imaginação que se tenha, a única coisa que se consegue observar e registar é que a característica “Departamento” assume resultados diferentes e atribuir-lhes um nome. Não conseguimos, pelos menos de forma consensual, definir se a ocorrência “Recursos Humanos” é melhor ou pior do que a de “Vendas”, sabemos apenas que são diferentes. Trata-se de uma situação típica de utilização de uma escala qualitativa nominal, que neste caso, a título de exemplo, poderia ser “Vendas”, “Produção”, “Marketing” e “Financeira”.

A escala qualitativa nominal apenas permite registar as diferentes formas que a uma determinada característica pode assumir atribuindo-lhes etiquetas (nomes). Como não existe uma lógica que permita ordenar, de forma consensual, as diferentes formas que a característica pode assumir, torna-se perfeitamente indiferente a ordem pela qual são dispostas na escala as etiquetas atribuídas a estas. Em termos de detalhe a escala qualitativa nominal encontra-se no nível mais baixo, uma vez que a informação registada se resume apenas à identificação do tipo de ocorrência da característica.

No âmbito do hipotético estudo referido anteriormente outras das informações que é necessário registar é o nível de desempenho dos colaboradores. Neste caso, quando se

comparam dois colaboradores, não só é possível identificar se têm desempenhos idênticos, como, caso não tenham, é possível perceber qual é que teve melhor desempenho. Assim, a escala de medida não só tem que permitir distinguir as diferentes ocorrências da característica desempenho – tal como a escala nominal –, como permitir ordenar as mesmas de acordo com o nível de qualidade do desempenho. Neste caso e a título de exemplo a escala podia ser “Baixo”, “Médio” e “Elevado”.

A escala qualitativa ordinal, face à nominal, acrescenta uma segunda camada de informação, que permite hierarquizar as diferentes ocorrências da característica, recorrendo a um critério lógico e aceite de forma consensual. Em termos do detalhe da informação registada estas escalas encontram-se um nível acima das escalas nominais.

Imaginando agora que o nível de desempenho dos colaboradores era outra das informações a recolher e que este é avaliado através da percentagem média de cumprimento dos seus objetivos individuais. Neste caso passa a ser possível exprimir o nível de desempenho de cada colaborador através de uma percentagem, isto é, numa escala que varia entre 0 e 100%. Se o colaborador A tiver 30% e o B 60%, não só ficamos a saber que têm níveis de desempenho diferentes e que o B teve um melhor desempenho, como sabemos quão melhor o B esteve face ao A. Sabemos que o nível de desempenho do B está 30 pontos percentuais acima do de A ou que foi o dobro deste. Esta camada de informação adicional, que consiste na quantificação da distância entre os diferentes níveis de desempenho, é o que nos permite considerar a escala de medida como quantitativa.

Os dados resultantes do registo de um fenómeno através de uma escala quantitativa distinguem-se facilmente dos dados recolhidos usando uma escala qualitativa, uma vez que são representados por valores, que traduzem a intensidade com que a características ocorre. Uma escala quantitativa, para além de permitir identificar ocorrências distintas e de as ordenar segundo um determinado critério lógico, também permite quantificar as diferenças entre estes. Existem dois tipos distintos de escalas quantitativas, que se designam por intervalares e de razão ou rácio, cuja única diferença reside na localização relativa do valor zero na escala.

As escalas intervalares caracterizam-se por serem escalas quantitativas em que a posição relativa do ponto zero é arbitrária, i.e., a posição relativa do valor zero na escala pode variar com as unidades de medida utilizadas. Um exemplo clássico do uso de uma escala intervalar é o registo de temperaturas, em que nas diferentes escalas de medida existentes (e.g., °C e °F) o ponto zero

tem uma posição relativa diferente. Como consequência desta arbitrariedade da posição relativa do ponto zero, se tivermos dois recipientes com água, um a 10°C e outro a 20°C, não se pode afirmar que o primeiro tem o dobro da temperatura do segundo. A razão de ser desta aparente contradição é a de que se passarmos os valores de °C para °F, passamos a ter os valores de 50°F e 68°F, para o primeiro e segundo recipiente, respetivamente. Como se pode constatar quando as temperaturas são expressas em °F a temperatura do segundo recipiente está longe de ser o dobro da do primeiro, o que se deve unicamente a uma posição relativa do ponto zero. Assim, numa escala intervalar os níveis são mutuamente exclusivos, têm uma ordem lógica e são escalonados de acordo com as características que possuem. As diferenças iguais nas características são representadas por diferenças iguais nas quantificações dos níveis e o valor zero é apenas mais um nível da escala. Em termos do detalhe da informação este tipo de escalas está claramente um nível acima de qualquer um dos tipos de escala qualitativa anteriormente referidos, uma vez que permite a quantificação das distâncias entre os diferentes níveis do grau de ocorrência da característica em análise.

As escalas de razão ou rácio são escalas quantitativas em que o ponto zero é absoluto, i.e., independentemente da unidade escolhida para exprimir os valores, a posição relativa do valor zero na escala é sempre a mesma. Um possível exemplo do uso de uma escala de razão ou rácio é a medição de uma distância. Se tomarmos como exemplo as distâncias entre Tomar e Lisboa (140 km) e entre Tomar e Beja (280 km), podemos com a toda a segurança afirmar que Beja está ao dobro da distância de Tomar do que está Lisboa. Esta afirmação é verdadeira, porque mesmo que as duas distâncias passem a ser expressas em milha, polegada, jarda ou qualquer outra unidade de medida, o valor da distância entre Tomar e Beja, será sempre o dobro do valor correspondente à distância entre Tomar e Lisboa. Os níveis são mutuamente exclusivos, têm uma ordem lógica e são escalonados de acordo com as características que possuem. As diferenças iguais nas características são representadas por diferenças iguais nas quantificações dos níveis e o valor zero reflete de facto a ausência da característica em causa. No que diz respeito ao detalhe com que a informação é registada as escalas de razão ou rácio encontram-se no nível máximo, permitindo a aplicação de qualquer operação matemática aos dados que se obtêm por recurso a este tipo de escala.

1.1. Análise Comparativa das Diferentes Escalas de Medida

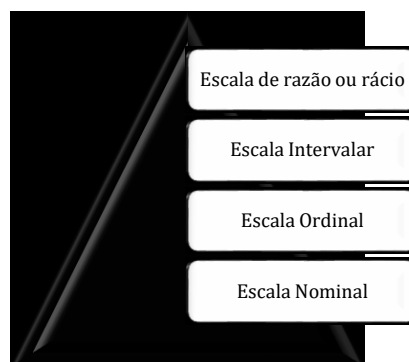
Assumindo um ponto de vista mais pragmático, a principal diferença que ressalta das descrições individuais das diferentes escalas de medida apresentadas anteriormente, reside na sua capacidade de reter informação, ou seja, no detalhe com que a ocorrência de uma determinada característica é registada (Hugh Coolican, 2014). Quanto maior for o detalhe no registo da informação maior será o potencial de informação passível de ser extraída no processo de análise dos dados. O que não quer dizer que se deve usar sempre a escala que regista a informação com maior detalhe (quantitativa de razão), uma vez que nem todas as características são compatíveis com este tipo de escala. Desta forma, torna-se evidente que a correta seleção da escala é um fator determinante para a qualidade dos resultados da análise. A natureza da escala de medida é determinada pela natureza da característica em análise e determina a natureza da informação recolhida.

Nos casos típicos, como os dos exemplos referidos anteriormente, é relativamente simples escolher a escala de medida mais adequada. O problema levanta-se nas situações fronteira que deverão ser alvo de uma análise mais cuidada, tendo sempre presente a hierarquia que existe entre as diferentes escalas no que diz respeito ao detalhe da informação recolhida e as suas consequências em termos da análise dos mesmos.

Existem algumas características, que embora sejam passíveis de ser quantificadas, são frequentemente registadas não em termos do seu valor absoluto, mas através de uma lista com alguns intervalos de valor. É o caso da informação sobre a idade ou sobre os rendimentos, que, apesar de quantificáveis, é frequentemente recolhida por intervalos de valores, o atribui uma natureza qualitativa aos dados. O investigador deixa de conhecer a idade ou o rendimento da pessoa, sabendo apenas em que intervalo de valores este se situa. Como nestes casos é possível estabelecer um critério para ordenar os diferentes intervalos está-se perante uma escala ordinal. Contudo, é importante que o utilizador tenha consciência que, sempre que recolhe dados quantificáveis sob a forma de intervalos, está a usar uma escala qualitativa ordinal, fazendo com que a informação recolhida, em termos de detalhe do registo, se encontre a um nível substancialmente inferior ao que estaria se fosse recolhida usando uma escala quantitativa, o que terá consequência no detalhe das conclusões extraídas.

Podemos assim concluir que existem quatro níveis de escala de medidas: escala nominal que categoriza sem ordenar, a escala ordinal que categoriza e ordena, a escala intervalar que quantifica, ordena e estabelece uma unidade igual na escala e finalmente a escala de razão ou rácio que quantifica, ordena, estabelece uma unidade igual na escala e o valor zero corresponde à inexistência da característica (Ramos Pinto, 2012).

Figura 1 – Comparação dos diferentes tipos de escala, em termos do detalhe da informação registrada



Fonte: Elaboração dos autores

1.2. Fiabilidade e Validade

A utilização de uma nova escala de medida deve ser sempre antecedida de uma análise da sua capacidade de medir corretamente a característica em causa, através da sua aplicação num ensaio piloto a uma amostra de dimensões mais reduzidas (Wimmer e Dominick, 2011), usualmente designada por pré-teste. Para que seja útil e credível uma escala de medida deve ser fiável e válida (Hugh Coolican, 2014). A fiabilidade lida com a consistência e estabilidade das medições, isto é, está relacionada com a precisão do método de medição, que deve produzir resultados idênticos quando repetido sucessivamente, nas mesmas condições e sobre a mesma unidade de análise. A validade, por sua vez, diz respeito à veracidade das medições e está relacionada com a capacidade de medir corretamente a ocorrência da característica em análise, o que é muito mais complicado de verificar do que no caso da fiabilidade (Wimmer e Dominick, 2011). A análise detalhada dos conceitos de fiabilidade e validade está fora do âmbito deste artigo, recomendando-se a leitura de Wimmer e Dominick (2011) e Coolican (2014).

2. A CReSAP: Medição do Mérito no Recrutamento e Seleção

Ao levar a cabo uma reflexão sobre o mérito e os desafios da sua medição, no contexto português, há pelo menos uma distinção que se afigura como natural: a distinção entre a medição do mérito no acesso aos cargos de direção superior da administração pública e a medição do mérito no desempenho de serviços, dirigentes e trabalhadores.

No que respeita ao mérito e à sua medição no recrutamento e seleção de candidatos para cargos de direção superior na administração pública, trata-se de uma competência atribuída à Comissão de Recrutamento e Seleção para a Administração Pública (CReSAP), instituída pela Lei n.º 64/2011, de 22 de dezembro (Portugal, 2011).

A CReSAP tem como fundamento teórico os desenvolvimentos que derivaram do trabalho pioneiro de Marston (1928). Nessa senda, a medição do mérito ocorre ao longo de dois eixos, formando quatro quadrantes que podem ser sintetizados na tabela 1:

Tabela 1 – Possibilidades de perfis de liderança

	Orientação para Tarefas	Orientação para Pessoas
Orientação para Ação	Domínio (Características associadas: determinação; assertividade; competitividade)	Influência (Características associadas: persuasão; intuição; comunicação)
Orientação para Reflexão	Conformidade (Características associadas: cautela; sistematização; objetividade)	Estabilidade (Características associadas: passividade; persistência; ponderação)

Fonte: adaptado da teoria DISC de Marston (1928).

É importante frisar que, relativamente à tabela anterior, os candidatos podem apresentar inúmeras combinações e gradações distintas, pelo que o papel desta agência independente do Estado consiste em fazer corresponder aos perfis definidos para os vários cargos os perfis dos vários candidatos aos mesmos¹. Por entendimento da CReSAP, a caracterização dos candidatos em termos dos seus perfis de liderança é complementada por critérios de avaliação curricular e de desempenho em entrevista.

¹ Neste sentido, o trabalho desenvolvido pela CReSAP poderá ser bem modelado pela teoria de leilões. Ao leitor que deseje iniciar o aprofundamento desta abordagem relativamente recente recomenda-se a leitura de Castelar *et al.* (2010).

Para além da procura de identificação, nos candidatos, de cinco traços de personalidade padrão (humildade, adaptação à mudança, sociabilidade, agradabilidade e equilíbrio emocional), a Comissão de Recrutamento e Seleção para a Administração Pública procede a uma subdivisão de competências para aferição do mérito em 12 critérios de gestão (cuja ponderação decorre de deliberação do júri de cada procedimento concursal em função do perfil pretendido e é divulgada publicamente no edital de cada concurso) e em 12 critérios comportamentais (decorrentes de aprovação de regulamento) que, quando combinados, permitem discriminar a qualidade relativa dos candidatos (tabela 2).

Tabela 2 – Critérios de gestão e critérios comportamentais

12 Critérios de Gestão	12 Critérios Comportamentais
Liderança	Criatividade
Colaboração	Determinação
Motivação	Individualidade
Orientação estratégica	Voluntariedade
Orientação para resultados	Confiabilidade
Orientação para o cidadão e o serviço público	Autoconfiança
Gestão da mudança e inovação	Paciência
Sensibilidade social	Reflexão e concentração
Aptidão	Persistência
Experiência profissional	Adaptabilidade
Formação académica	Perfeccionismo
Formação complementar	Sensibilidade e perspicácia

Fonte: CReSAP

Estes critérios são avaliados, por um lado, à luz da informação disponibilizada aquando da resposta, por parte dos candidatos, ao Questionário de Autoavaliação Curricular, constante do Regulamento de Tramitação dos Procedimentos de Recrutamento e Seleção dos Cargos de Direção Superior na Administração Pública (disponível em <http://www.cresap.pt>) e, por outro lado, à luz da informação constante do currículo dos mesmos.

Um importante desafio na mediação do mérito dos candidatos a cargos de direção superior na administração pública surge quando se procura impossibilitar que o júri de cada concurso possa determinar arbitrariamente, por intermédio da definição de ponderações *made to measure* dos critérios (subvalorizando ou sobrevalorizando determinados critérios), o resultado final. A resposta a este desafio surgiu por intermédio da obrigatoriedade de os júris respeitarem

limites máximos e mínimos, previamente fixados, para as ponderações de grupos de critérios (tabela 3)².

Tabela 3 – Limites máximos e mínimos para ponderações

Grupo de critérios	Ponderação Mínima	Ponderação Máxima
<ul style="list-style-type: none">▪ Liderança▪ Colaboração▪ Motivação	15%	30%
<ul style="list-style-type: none">▪ Orientação estratégica▪ Orientação para resultados▪ Gestão da mudança e inovação	15%	30%
<ul style="list-style-type: none">▪ Orientação para o cidadão e o serviço público▪ Sensibilidade social	15%	20%
<ul style="list-style-type: none">▪ Aptidão	5%	15%
<ul style="list-style-type: none">▪ Experiência profissional▪ Formação acadêmica▪ Formação complementar	15%	30%

Fonte: CReSAP

Desta forma, torna-se possível à CReSAP definir um conjunto de perfis-padrão para os candidatos, necessariamente dependentes de uma cuidadosa análise dos requisitos definidos pelas tutelas, nomeadamente, as áreas de formação e especialização preferenciais, as principais funções e responsabilidade associadas ao cargo de direção a concurso, as características do mandato de gestão e a identificação do cargo a concurso e das competências que lhe estão associadas. Desta forma, é possível alterar a tabela 1 de modo a associá-la aos principais perfis-padrão (tabela 4):

Tabela 4 – Possibilidades de perfis de liderança e perfis-padrão sugeridos

	Orientação para Tarefas	Orientação para Pessoas
Orientação para Ação	Domínio Dirigente Superior de 1.º grau	Influência Dirigente Superior de 2.º grau (Perfil de Gestor)
Orientação para Reflexão	Conformidade Dirigente Superior de 2.º grau (Perfil Técnico)	Estabilidade Dirigente Superior de 2.º grau (Perfil de Jurista)

Fonte: adaptado da teoria DISC de Marston (1928) e CReSAP.

² Note-se que, desde que observadas estas limitações, o júri é soberano para deliberar acerca da melhor combinação de ponderações para o perfil desejável no concurso relativo a um determinado cargo.

A caracterização destes quatro perfis-padrão pode ser aprofundada por intermédio da construção de quadros comparativos de perfis. Uma possível solução para o quadro comparativo da importância dos critérios comportamentais dos perfis-padrão pode ser encontrada na tabela 5:

Tabela 5 – Caracterização dos perfis-padrão ao nível das competências comportamentais

	Dirigente Superior de 1.º grau	Dirigente Superior de 2.º grau		
Criatividade	Intermédia	Intermédia	Alta	Intermédia
Determinação	Alta	Alta	Alta	Alta
Individualidade	Intermédia	Baixa	Intermédia	Baixa
Voluntariedade	Intermédia	Alta	Baixa	Baixa
Confiabilidade	Alta	Alta	Intermédia	Baixa
Autoconfiança	Intermédia	Alta	Baixa	Baixa
Paciência	Baixa	Baixa	Baixa	Intermédia
Reflexão e concentração	Intermédia	Intermédia	Intermédia	Intermédia
Persistência	Baixa	Baixa	Baixa	Alta
Adaptabilidade	Baixa	Baixa	Alta	Intermédia
Perfeccionismo	Baixa	Baixa	Alta	Alta
Sensibilidade e perspicácia	Alta	Alta	Alta	Intermédia

Fonte: adaptado de CReSAP.

Por outro lado, uma possível solução para o quadro comparativo da importância dos critérios de gestão dos perfis-padrão pode ser encontrada na tabela 6:

Tabela 6 – Caracterização dos perfis-padrão ao nível das competências de gestão

	Dirigente Superior de 1.º grau	Dirigente Superior de 2.º grau (Perfil de Gestor)	Dirigente Superior de 2.º grau (Perfil Técnico)	Dirigente Superior de 2.º grau (Perfil de Jurista)
Liderança	Alta	Intermédia	Alta	Alta
Colaboração	Intermédia	Alta	Baixa	Intermédia
Motivação	Baixa	Baixa	Baixa	Intermédia
Orientação estratégica	Alta	Intermédia	Intermédia	Baixa
Orientação para resultados	Alta	Alta	Alta	Alta
Orientação para o cidadão e o serviço público	Intermédia	Intermédia	Baixa	Alta
Gestão da mudança e inovação	Baixa	Intermédia	Intermédia	Baixa
Sensibilidade social	Baixa	Baixa	Baixa	Intermédia
Aptidão	Intermédia	Baixa	Baixa	Baixa
Experiência profissional	Alta	Alta	Alta	Alta
Formação académica	Intermédia	Alta	Alta	Intermédia
Formação complementar	Baixa	Intermédia	Intermédia	Intermédia

Fonte: adaptado de CReSAP.

Nestes termos, o desafio último de medição do mérito no recrutamento e seleção de candidatos para cargos de direção superior na administração pública, traduzido pela definição das ponderações a atribuir aos critérios constantes da tabela 6, poderia ser resolvido por intermédio dos valores constantes na tabela 7:

Tabela 7 – Ponderações propostas para os perfis-padrão ao nível das competências de gestão (em percentagem)

	Dirigente Superior de 1.º grau	Dirigente Superior de 2.º grau (Perfil de Gestor)	Dirigente Superior de 2.º grau (Perfil Técnico)	Dirigente Superior de 2.º grau (Perfil de Jurista)
Liderança	10%	8%	10%	10%
Colaboração	8%	10%	7%	8%
Motivação	7%	6%	7%	7%
Orientação estratégica	10%	8%	8%	7%
Orientação para resultados	10%	10%	10%	10%
Orientação para o cidadão e o serviço público	8%	8%	8%	10%
Gestão da mudança e inovação	7%	8%	8%	7%
Sensibilidade social	7%	7%	7%	8%
Aptidão	8%	7%	7%	7%
Experiência profissional	10%	10%	10%	10%
Formação académica	8%	10%	10%	8%
Formação complementar	7%	8%	8%	8%

Fonte: adaptado de CRE SAP.

Ao leitor interessado no aprofundamento da temática do mérito na administração pública, sugere-se a consulta de Bilhim (2012, 2013), Bilhim e Garcia (2013) ou Bilhim e Correia (2016).

3. A Medição do Mérito no Desempenho

Na administração pública contemporânea, tem sido notório um maior enfoque na medição do desempenho no governo, nas organizações públicas e, consequentemente, nos seus funcionários (Vandenabeele e Hondegheem, 2008:243). Verifica-se, na literatura, um crescente enfoque no debate sobre a predominância, nas agências do setor público, da mentalidade gestonária associada à Nova Gestão Pública e que enfatiza a eficiência, a efetividade e a qualidade e visa tornar estas organizações mais orientadas para os resultados (Rodwell, Noblet, e Allisey, 2011). A avaliação de desempenho é vista como um instrumento de gestão fundamental no contexto da Nova Gestão Pública (Gianakis, 2002), que promove a *accountability*, ao fornecer

informação sobre a eficiência e eficácia da administração dos serviços, e é considerada como um dos principais motores da reforma do sector público nos últimos anos (Greiling, 2005).

A avaliação de desempenho foi introduzida na administração pública portuguesa em 2007 com a criação do Sistema Integrado de Avaliação de Desempenho da Administração Pública (SIADAP) (Portugal, 2007). Este sistema abrange a totalidade dos funcionários públicos da administração pública, desde os níveis mais baixos até aos dirigentes, com exceção das carreiras especiais, que têm sistemas de avaliação próprios. Neste artigo começaremos por analisar o SIADAP e de seguida analisaremos a avaliação de desempenho em dois setores relevantes da sociedade, o Ensino Superior e a Justiça, nos quais existem duas das mais expressivas carreiras especiais, Professores Universitários e os Magistrados.

3.1. O Sistema Integrado de Avaliação de Desempenho da Administração Pública (SIADAP)

Relativamente ao SIADAP e às conceções que o sustentam, torna-se imperativo proceder a uma breve resenha.

O SIADAP, em confronto com o anterior paradigma de utilização de normas e regulamentos, configura um instrumento técnico que se pretendeu de cariz inovador, procurando proporcionar ao gestor público um conjunto de ferramentas que possibilitassem mudanças na forma como o mérito é aferido no seio da máquina administrativa, pressupondo:

- Uma definição clara de objetivos e a necessária comunicação à organização;
- A atribuição de recursos em ordem à consecução dos objetivos estabelecidos;
- O controlo de custos e, paralelamente, o estímulo motivacional dos colaboradores;
- Convergência simultânea para o incremento da eficácia, eficiência e qualidade, consonante com uma visão estratégica de pró-atividade.

Com a oficialização do SIADAP, na sua mais recente formulação, promulgado pela Lei n.º 66-B/2007 (Portugal, 2007)³, são incluídos conceitos refletidos na própria designação, tais como sistema integrado, gestão, avaliação e desempenho. Estes conceitos são associados a três

³ Entretanto alterada pela Lei n.º 64-A/2008, de 31 de dezembro (Portugal, 2008a), pela Lei n.º 55-A/2010, de 31 de dezembro (Portugal, 2010) e pela Lei n.º 66-B/2012, de 31 de dezembro (Portugal, 2012).

níveis de medicação de mérito no desempenho: o primeiro nível, correspondente ao SIADAP 1, que trata da avaliação do desempenho dos serviços da administração pública (como por exemplo Direções-Gerais ou Institutos Públicos); o segundo nível, correspondente ao SIADAP 2, que trata da avaliação do desempenho dos dirigentes da administração pública (dirigentes superiores e intermédios); e o terceiro nível, correspondente ao SIADAP 3, que trata da avaliação do desempenho dos trabalhadores da administração pública (numa ótica já muito focada na gestão de recursos humanos). Por isso mesmo, suscita particular interesse conhecer a interpretação da lei e a intenção do Legislador.

Os princípios-chave, nos quais o SIADAP se baseia, devem ser entendidos como um meio de assegurar a sua articulação com o ciclo de gestão das várias entidades, de forma coerente e integrada. Conceptualmente também se procurou estabelecer a ligação entre este sistema de avaliação e o mérito, ao responsabilizar os intervenientes pelos êxitos e fracassos organizacionais, uma vez que esses mesmos intervenientes são agentes ativos associados ao desempenho da organização.

Procurou-se ainda, respondendo à necessidade de salvaguardar a universalidade de aplicação do sistema, acautelar princípios que lhe conferissem um carácter manifestamente universal, numa tentativa de expandir a sua aplicabilidade a organizações de natureza e função muito diferenciadas, e ao mesmo tempo permitindo acomodar alguma flexibilidade em termos de adaptação a contingências inesperadas.

A estruturação do SIADAP, tal como acontece com qualquer sistema credível de avaliação, foi efetuada com o objetivo prevalecente de transparência e imparcialidade nos métodos de avaliação propostos, atestando a comparabilidade dos resultados obtidos (Correia, 2012). Este modelo de avaliação de desempenho também se alicerçou em três grandes vetores-chave, associados a eficácia, eficiência e qualidade. Na verdade, o objetivo último da conceção deste sistema antecipa o envolvimento ativo de todos os participantes empenhados no processo avaliativo, designadamente, dirigentes, colaboradores e utilizadores dos serviços prestados pela administração pública (Correia 2011, 2012), para uma adequada divulgação de propósitos e resultados quer nas organizações públicas, quer na sociedade em geral.

Durante a existência (ainda relativamente curta) do SIADAP, nesta sua mais recente roupagem, foi possível constatar, do ponto de vista da sua aplicação prática, que este sistema de avaliação dos serviços almeja uma estreita articulação com o modelo de planificação de cada

ministério, em parte, como instrumento de avaliação da observância dos objetivos estratégicos plurianuais e, genericamente, como instrumento focado nos objetivos anuais e nos planos de atividades da administração pública portuguesa. Idealmente, o processo de gestão integrada do desempenho deveria acompanhar o ciclo de gestão, que previamente calendarizaria e fixaria em termos quantitativos os objetivos de desempenho das atividades e programas, sendo o desempenho mensurado e reportado no final. Associada a esta questão da definição de objetivos, quantitativos e calendarizados, referentes ao ano seguinte, está a sua relação intrínseca com o grau de complexidade da gestão do serviço em causa, englobando assim, mais ou menos extensivamente, a missão inerente ao serviço, as suas atribuições específicas, a identificação plurianual dos objetivos estratégicos estabelecidos superiormente, os compromissos assumidos pelas chefias na carta de missão, os resultados da avaliação do desempenho, e as disponibilidades orçamentais e respetivos mapas de pessoal. É de realçar que, na maioria dos casos, se assiste a uma articulação com o plano de atividades e que a regularidade de monitorização aliada à revisão de objetivos, se têm processado sem que haja registo de falhas graves ou de verdadeiros obstáculos à sua concretização (Correia, 2012). Tal como previsto na lei, os resultados alcançados pelos serviços também constam do relatório de atividades. Face a este novo conjunto de atividades, a articulação com o sistema de planeamento envolve uma coordenação constante entre os diversos serviços e as entidades que, a nível ministerial, se encarregam de matérias de planeamento, estratégia e avaliação, de que são exemplo, os Gabinetes de Planeamento, Estratégia, Avaliação e Relações Internacionais (GPEARI). Pelas razões acima expostas, e apesar de ser questionável o seu grau de envolvimento, importa reter que o SIADAP apresenta claramente um cariz inspirado nos princípios da Nova Gestão Pública, enquanto agente libertador de uma política organizativa que alia à gestão a aplicação de regulamentos e regras, tendo a particularidade de procurar constituir uma forma diferente de encarar os problemas que as organizações enfrentam, não compaginável com a anterior, e cada vez mais enraizado na prática e não tanto em termos maioritariamente normativos. Sob um outro ângulo, e considerando a escassez generalizada de recursos, quer materiais, quer financeiros, ou humanos e temporais, pode depreender-se que o SIADAP teve como propósito materializar um conjunto de critérios objetivos, proporcionando um meio de aferição de problemas suscetíveis de resolução, bem como dos que se afiguram como irresolúveis, dadas as contingências do mundo real. Este posicionamento introduz no espírito do SIADAP uma alteração no sentido do (novo)

paradigma organizacional da Nova gestão Pública, e preconiza a utilização de instrumentos capazes de aferir o padrão de eficácia, eficiência e qualidade nessa mesma gestão, um marco fundamental da evolução do modelo de avaliação, que recupera o propósito de transparência de atuação e de prestação de contas da Administração Pública e, acima de tudo, visa a melhoria do exercício das suas atribuições perante os cidadãos, em geral.

Consequentemente, não surpreende que se deva enfatizar a importância deste novo paradigma instituído pelo SIADAP, centrando as atenções nos três grandes pilares de avaliação: eficácia, eficiência e qualidade.

É a própria Lei n.º 66-B/2007 que se refere ao conceito de eficácia como a “*medida em que um serviço atinge os seus objetivos e obtém ou ultrapassa os resultados esperados*” (Portugal, 2007). Neste âmbito, admite-se que a origem da gestão assenta precisamente na persecução da resolução de questões de ineficácia, entendendo-se a eficácia como o primeiro degrau na escala ascendente da gestão (Correia, 2011, 2012). Neste sentido, a dimensão eficácia invocada no SIADAP pretende computar este alicerce primordial da gestão.

Por outro lado, a eficiência pode denotar a ambição de fazer mais, quantitativamente ou qualitativamente, com a mesma disponibilidade de recursos, ou além disso a capacidade de realizar uma quantidade análoga de trabalho com menor dotação de recursos. Este pilar, fortemente relacionado com o fator temporal e com o cumprimento de prazos e a sua superação, constitui o que pode ser tratado como o segundo patamar na escala progressiva de aprimoramento da gestão. O facto de constar entre os três critérios de avaliação de desempenho dos serviços demonstra a sua relevância.

Em contrapartida, o pilar qualidade pretende agregar ao processo avaliativo dos serviços *inputs* externos e, ao mesmo tempo, nortear os processos internos numa lógica focada em servir os cidadãos, correspondendo ao terceiro patamar na escala de aprimoramento gradual da gestão. A Lei n.º 66-B/07 faz referência aos objetivos de qualidade como o “*conjunto de propriedades e características de bens ou serviços, que lhes conferem aptidão para satisfazer necessidades explícitas ou implícitas dos utilizadores*” (Portugal, 2007), assinalando a sua especial importância.

Em síntese, apesar do tecnicismo e da complexidade legislativa, pode afirmar-se que a globalidade dos objetivos do SIADAP se resumem a um conjunto elementar de ideias-chave. Efetivamente, o SIADAP deixa trespassar para a administração pública o propósito de obtenção de um aperfeiçoamento mediante a consolidação de métodos de avaliação objetivos que

possibilitam diferentes intervenções: reconhecer debilidades no campo formativo, desenvolver o estímulo motivacional entre os colaboradores, aperfeiçoar a estrutura de procedimentos, melhorar a comunicação tanto a nível interno como externo, com a preocupação de promover total transparência e, além disso, apoiar os métodos de decisão estratégica da organização (Correia, 2012). O culminar de todas estas atuações traduz-se na possibilidade de deteção intencional e sistemática de dissemelhanças de desempenho, na execução das funções de serviço público, em serviços e pessoas.

No contexto do SIADAP, deve-se também fazer referência à ideia da necessidade de assegurar o cumprimento daquilo a que se pode dar o nome de fatores críticos de sucesso e que incluem:

- A premência de garantir que a perceção do conjunto dos participantes relativamente ao sistema de avaliação não o toma apenas como uma técnica mas antes como um novo modelo de gestão pública.
- O bom senso de permitir que o SIADAP obedeça a um calendário de desenvolvimento gradual, objeto de permanente reajustamento e sem sucumbir à tentação do desejo de alcançar soluções perfeitas, já que o erro constitui, necessariamente, um elemento de aprendizagem.

Se se praticarem estes princípios, o SIADAP poderá transformar-se, a médio ou longo prazo, num instrumento mais abrangente, de aferição de mérito no desempenho, de dotação laboral, de constante ascensão qualitativa do serviço, do dirigente e do trabalhador, por intermédio: da conceção de padrões de referência entre serviços análogos; e da seleção de lideranças idóneas, aptas a promover a integração e crescimento dos trabalhadores e da própria sociedade no sentido de propiciar às instituições públicas o prosseguimento das incumbências inerentes.

3.2. O Caso de Estudo dos Desafios da Medição do Mérito no Desempenho no Setor do Ensino Superior

O Ensino Superior, desde a sua criação no século XII na Europa, tem assumido um papel central na sociedade, sendo frequentemente chamada a desempenhar funções essenciais no seu contexto social (Altbach, 2008). Atualmente, estas Instituições têm uma intervenção

determinante no desenvolvimento humano, social e económico. A tendência é para que aumente o envolvimento com as comunidades em que estão envolvidas, assumindo-se como atores promotores da mudança social (Escrigas, 2014).

“A avaliação do desempenho docente em Instituições de Ensino Superior (IES), de forma similar à avaliação do desempenho de qualquer profissional nos diversos campos de atuação e nos diversos sistemas e processos produtivos, é de fundamental importância para o desenvolvimento profissional dos docentes e para o crescimento institucional, ajudando a garantir que as metas qualitativas e quantitativas esperadas pela sociedade possam ser alcançadas”. (Embiruçu, Fontes e Almeida, 2010:797).

No caso particular da realidade portuguesa torna-se interessante analisar a avaliação de desempenho no Ensino Superior, uma vez que: 1) Em 2005, o setor inicia uma profunda reestruturação da sua oferta educativa, que decorre da aplicação do Acordo de Bolonha; 2) O regime jurídico das instituições de ensino superior é também profundamente alterado em 2007, adequando a sua estrutura e funcionamento ao modelo gestor; 3) Em 2009 são introduzidas profundas alterações nas carreiras docentes, que obrigam a um esforço de qualificação dos recursos; 4) O sistema de avaliação de desempenho também é profundamente alterado em 2009.

3.2.1. O Setor do Ensino Superior em Portugal

A rede de Ensino Superior divide-se em dois subsistemas: um de natureza pública e em cujo modelo de financiamento, embora inclua receitas próprias, assenta em grande parte nas transferências do orçamento de Estado; e um de natureza privada em que o modelo de financiamento, embora possa incluir algum financiamento público, assenta essencialmente na capacidade de arrecadação de receita própria. Cada um destes subsistemas subdivide-se em dois tipos de ensino: o Ensino Superior Universitário, que tem como missão a formação avançada de natureza mais científica e de investigação, e confere os graus de licenciado, mestre e doutor; e o Ensino Superior Politécnico, que visa proporcionar formação avançada de natureza mais especializada, técnica e profissional, e confere os graus académicos de licenciado e de mestre, e o diploma de técnico superior profissional.

O subsistema público é composto por 13 Universidades, um Instituto Universitário, quatro Instituições Universitárias Militares e uma Instituição Universitária Policial.

Complementarmente, existe uma rede composta por 15 Politécnicos e cinco Escolas Superiores não integradas.

No caso do subsistema privado existem 10 Universidades, cinco Institutos Universitários e 11 Escolas Universitárias não integradas; complementadas com 67 Escolas Politécnicas.

A rede de Ensino Superior, em 2015, acomodou 349658 alunos e 32580 docentes, produziu 76892 diplomados e representou um investimento público de 3,8% do PIB. Num espaço de 10 anos a rede de Ensino Superior foi responsável por um quase duplicação do nível de escolaridade superior na população entre os 15 e os 64 anos, que passou de 11,10% em 2005 para 20,70% em 2015.

Tabela 8 – Quadro comparativo do número de docentes, alunos e diplomados no sistema de Ensino Superior em Portugal, entre 2005 e 2015

		2005			2015			Variação do Total
		Universidades	Politécnicos	TOTAL	Universidades	Politécnicos	TOTAL	
Docentes	Instituições Públicas	15.506	10.904	26.410	15.704	9.438	25.142	-4,80%
	Instituições Privadas	7.195	4.459	11.654	4.779	2.659	7.438	-36,18%
Alunos	Instituições Públicas	173.897	108.376	282.273	191.707	100.652	292.359	3,57%
	Instituições Privadas	67.157	31.507	98.664	42.666	14.633	57.299	-41,93%
Diplomados	Instituições Públicas	25.283	23.901	49.184	42.389	20.578	62.967	28,02%
	Instituições Privadas	11.172	9.631	20.803	9.978	3.947	13.925	-33,06%

Fonte: Pordata.

As instituições de Ensino Superior Público, embora estejam sob a tutela do Ministério da Ciência, Tecnologia e Ensino Superior e recebam transferências do Orçamento de Estado, têm autonomia administrativa e financeira, que lhes permite gerar receitas próprias e ter alguma autonomia de gestão. Analisando os dados da tabela 8, verifica-se que, em 2015, o subsistema público acomodou 84% dos alunos e 77% dos docentes e produziu 82% dos diplomados, sendo cerca de dois terços destes números assegurados pelas Universidades Públicas.

3.2.2. Avaliação de Desempenho no Ensino Superior em Portugal

Nas Instituições de Ensino Superior portuguesas existem as carreiras de docente universitário, de investigador, de docente do ensino superior politécnico e a carreira do pessoal

não docente. A última, à semelhança dos restantes funcionários públicos, é avaliada através do SIADAP (descrito anteriormente), mas as restantes, por serem consideradas carreiras especiais, têm sistemas de avaliação específicos.

A título de exemplo iremos analisar o sistema de avaliação de desempenho dos docentes universitários vinculados a Instituições Públicas, uma vez que estes, em 2015, totalizavam 15704 e representavam 48% dos docentes que integravam a rede de ensino superior em Portugal.

De acordo com o artigo 4.º do Estatuto da Carreira Docente Universitária (ECDU) “cumpre, em geral, aos docentes universitários as seguintes funções: 1) Realizar atividades de investigação científica, de criação cultural ou de desenvolvimento tecnológico; 2) Prestar o serviço docente que lhes for distribuído e acompanhar e orientar os estudantes; 3) Participar em tarefas de extensão universitária, de divulgação científica e de valorização económica e social do conhecimento; 4) Participar na gestão das respetivas instituições universitárias; 5) Participar em outras tarefas distribuídas pelos órgãos de gestão competentes e que se incluam no âmbito da atividade de docente universitário” (Portugal, 2009).

A carreira prevê, de acordo com os artigos 2.º e 5.º do ECDU, três categorias: Professor Catedrático, a que estão reservadas funções de coordenação e gestão das dimensões científica, pedagógica e de investigação; Professor Associado, a quem estão atribuídas funções de coordenação intermédia e de coadjuvação dos professores catedráticos; Professor Auxiliar, que corresponde à entrada na carreira académica e tem como funções primordiais a docência e a investigação. O acesso ao primeiro nível da carreira (professor auxiliar) e a transição para as restantes categorias, designada por promoção, ocorre obrigatoriamente por concurso público internacional e no caso de professor catedrático, implica que o candidato seja detentor do título académico de agregado. Em cada uma destas categorias existem 4 escalões remuneratórios, cuja progressão, até 2009, era automática a cada três anos e, após esta data, passou a depender de um sistema de créditos, obtidos através da avaliação do desempenho de cada docente no final de cada triénio.

O ECDU, no seu artigo 74.º-A, estipula que o regime de avaliação de desempenho dos docentes universitários está subordinado a um conjunto de 14 princípios, que, de forma resumida, visam garantir a universalidade, obrigatoriedade, *accountability*, imparcialidade, coerência, flexibilidade e previsibilidade do processo. Um dos princípios vincula o processo à consideração das cinco categorias de funções genericamente associadas ao desempenho da

função de docente universitário, estipuladas no artigo 4.º do ECDU e enunciadas anteriormente. Analisando os regulamentos de avaliação de desempenho dos docentes das principais universidades portuguesas, Lisboa, Porto, Nova de Lisboa, Coimbra e Aveiro, verifica-se que existe uma convergência em quatro vertentes de avaliação: 1) Ensino; 2) Investigação; 3) Extensão universitárias (atividades científicas e/ou culturais, que promovam a ligação e transferência de conhecimento para a sociedade); 4) Participação na gestão das universidades e suas unidades orgânicas. Esta tipificação da atividade dos docentes universitários encontra, de alguma forma, correspondência no trabalho de Boyer (1990), que tipificava a atividade docente em: 1) ensino; 2) investigação; 3) serviço profissional; 4) trabalho integrativo, e no trabalho de Braskamp e Ory (1994), que também assumia quatro vertentes na atividade docente, a saber: 1) ensino; 2) investigação e atividade criativa; 3) prática e serviço profissional; 4) cidadania e a participação cívica.

Cada uma das vertentes definidas é depois decomposta em subvertentes e indicadores, passíveis de serem medidos, de forma a permitir a avaliação do desempenho dos docentes. A forma de decomposição das vertentes (tipo e número de níveis/parâmetros considerados) embora dependa, inevitavelmente, da cultura organizacional das instituições, da sua área de atuação e do seu planeamento estratégico, tende a seguir os padrões internacionais do que se consideram serem os outputs desejáveis em cada uma das vertentes consideradas, principalmente no que diz respeito à investigação.

O grande desafio da avaliação do desempenho da atividade dos docentes universitários, resulta da grande diversidade de funções que potencialmente desempenham e da grande heterogeneidade que pode existir na distribuição do tempo que dedicam a cada uma delas. O que historicamente tem sido agravado pela maior dificuldade que tem existido em avaliar a atividade de ensino face às restantes, nomeadamente face à investigação, que, por apresentar outputs mais fáceis de medir e avaliar, tem assumido um papel preponderante nos sistemas de avaliação de desempenho. Já em 1990 Boyer (1990) defendia que, apesar de grande parte das universidades apregoarem a trilogia de ensino, pesquisa e serviço, no que diz respeito à atividade dos docentes universitários, quando se trata de fazer julgamentos sobre o desempenho profissional, raramente era atribuído igual mérito às três dimensões. O autor também defendia que o excessivo peso dado à investigação, gerava uma excessiva preocupação por parte dos docentes com esta faceta da sua atividade, o que a médio prazo se constituía como um obstáculo

à qualidade do ensino. Marsh and Roche (1997) também acentuaram a multidimensionalidade da atividade docente e, através de inquéritos aplicados a mais de um milhão de alunos, apuraram nove possíveis parâmetros de avaliação da vertente ensino, a saber: 1) valor da aprendizagem; 2) entusiasmo do docente; 3) organização e clareza; 4) interação em grupo; 5) relacionamento individual; 6) amplitude de cobertura; 7) avaliações/classificações; 8) trabalhos/leituras; 9) carga de trabalho/dificuldade. O desafio mantém-se, assim como a discussão sobre qual a melhor forma de medir o desempenho dos docentes, equilibrando as diferentes vertentes e acomodando o facto de que a proporção de dedicação a cada uma delas varia entre docentes e pode mesmo variar para o mesmo docente em momentos distintos da sua carreira. Theall (2010) considera que, atualmente, para que um docente possa ser considerado eficiente, não basta que este seja reconhecido pelos pares como um especialista na área que leciona (competência científica), mas terá que demonstrar que os seus alunos adquirem as competências. A responsabilidade pela aprendizagem foi parcialmente transferida dos alunos para os docentes e para as instituições, o que corresponde a uma alteração significativa no paradigma do ensino universitário, que não pode deixar de ser acomodada na avaliação de desempenho.

Para além da questão da definição dos indicadores a medir e da definição do seu peso relativo, existe um outro grande desafio, talvez ainda maior que os anteriores, que consiste no estabelecimento de metas que correspondam a um desempenho pleno do docente são desafios e onde a influência do contexto - onde se realça a cultura organizacional e a área de intervenção das instituições - é mais notória. Basta pensar-se na vertente da investigação, onde, inevitavelmente, se tem em consideração a produção científica do autor e onde, aparentemente, seria mais fácil estipular metas que traduzam o que é expectável um docente produzir. Contudo, se formos analisar o histórico e os padrões internacionais, a realidade - medida em termos do tipo e quantidade de outputs - na área da engenharia é completamente distinta da área das ciências sociais e, dentro desta última, entre a sociologia e as relações internacionais, só para citar alguns exemplos. Poder-se-á argumentar então que basta definir metas em função das especificidades de cada área e subárea, mas depois levanta-se o problema de saber como diferenciar as metas sem incorrer no risco de estar a favorecer áreas em detrimento de outras. Se olharmos para a participação em projetos de investigação, também não é fácil acomodar nas metas, por exemplo, a influência da política científica das entidades financiadores da

investigação, que claramente privilegiam o financiamento das áreas que consideram estratégicas, o que cria uma flagrante desigualdade entre os docentes em função da área onde investigam.

Voltando aos sistemas de avaliação de desempenho em vigor nas principais universidades portuguesas, a principal diferença que se encontra não está relacionada com o tipo de parâmetros que são usados, mas sim com a forma como são medidos e utilizados na avaliação do desempenho. As que assumem uma abordagem mais positivista tendem a usar modelos mais rígidos e quantificados, enquanto outras instituições, assumindo uma abordagem mais pós-positivista, tendem a valorizar mais a influência e especificidades do contexto, adotando modelos mais flexíveis em que o avaliador tem um pouco mais de autonomia e não se limita a contabilizar outputs e a comparar os resultados com as metas estabelecidas em regulamento.

Existem também experiência em outros países, como por exemplo Inglaterra, de indexação, ainda que parcial, das remunerações à avaliação do desempenho. Segundo Dolton, McIntosh e Chevalier (2003), tais medidas têm levado os professores concentrarem-se nos aspectos que são valorizados na avaliação de desempenho, o que pode ser prejudicial para a qualidade do ensino, principalmente, que for dado um peso excessivo à vertente investigação.

Em suma, existe um relativo consenso entre os especialistas que a avaliação de desempenho dos docentes universitários é um importante instrumento de gestão, melhorando a *accountability* e tendo um efeito sistêmico de melhoria da qualidade do ensino e do desempenho global das instituições (Boyer, 1990; Ramsden, 1991; Braskamp e Ory, 1994; Ramos e Moraes, 2000; Embiruçu, Fontes e Almeida, 2010). Contudo, os especialistas e a prática nas instituições universitárias portuguesas divergem na forma de a implementar, nomeadamente, na rigidez dos sistemas de avaliação e, conseqüentemente, na autonomia de avaliação que é concedida aos avaliadores.

3.3. O Caso de Estudo dos Desafios da Medição do Mérito no Desempenho no Setor da Justiça

Atualmente, face à progressiva subordinação da sociedade atual aos trâmites legais, o setor da justiça adquire uma relevância aumentada, abarcando a generalidade dos aspectos que compõem o quotidiano de indivíduos e instituições, e manifestando-se incrementalmente de forma transfronteiriça e mesmo global. É este o primeiro motivo que conduziu à seleção deste setor para apresentação de um estudo de caso: uma área de atuação do Estado eminentemente

abrangente, focada na prossecução das atribuições que lhe são intrínsecas, num contexto social de exigência incessante de aumento de eficácia, eficiência e qualidade. Este ponto é tão mais evidente quanto se tem em conta o enquadramento histórico recente de Portugal (nomeadamente a crise económica e financeira), o interesse da área científica da administração pública no estudo do sistema judicial português, e o seu elevado perfil e exposição mediática. O segundo motivo, prende-se com o facto de o setor da justiça ser um setor que, fruto da obrigatoriedade de separação de poderes associada às magistraturas, apresenta grupos de indivíduos que asseguram funções públicas sujeitos a regimes de aferição de mérito distintos.

O Ministério da Justiça, em Portugal, ilustra um caso paradigmático de uma organização pública. Enquanto departamento do Governo português, o Ministério da Justiça assegura o funcionamento do Sistema Judicial Português, sendo responsável pelo desenho, implementação e monitorização das políticas de Justiça, uma incumbência em que são salvaguardadas as ligações com os tribunais e o Ministério Público, o Conselho Superior da Magistratura e o Conselho Superior dos Tribunais Administrativos e Fiscais (Correia, 2012). Com restrições impostas ao poder executivo, pela autonomia do poder judicial, o Ministério da Justiça, diferentemente de outros ministérios, cinge-se, na generalidade das situações, a servir o sistema judicial em domínios de âmbito financeiro, legislativo e administrativo, competindo ao poder judicial a operacionalidade dos atos e tarefas atribuídos às suas instituições⁴.

Atualmente encontram-se tutelados pelo Ministério da Justiça 10 serviços da administração direta e indireta do Estado: Secretaria-Geral do Ministério da Justiça (SGMJ), Direção-Geral da Política de Justiça (DGPJ), Direção-Geral da Administração da Justiça (DGAJ), Direção-Geral de Reinserção e Serviços Prisionais (DGRSP), Inspeção-Geral dos Serviços de Justiça (IGSJ), Polícia Judiciária (PJ), Instituto de Gestão Financeira e Equipamentos da Justiça (IGFEJ), Instituto dos Registos e do Notariado (IRN), Instituto Nacional de Medicina Legal e Ciências Forenses (INMLCF) e, por fim, Instituto Nacional da Propriedade Industrial (INPI) (Correia e Moreira, 2016)⁵. Globalmente, este departamento do Estado acolhe perto de

⁴ Veja-se o exemplo da Polícia Judiciária, coordenada operacionalmente pelo Ministério Público mas dependente financeira, administrativa e organicamente do Ministério da Justiça.

⁵ Para além destes serviços, integram ainda o Ministério da Justiça o Gabinete da Ministra da Justiça, o Gabinete da Secretária de Estado Adjunta e da Justiça, o Gabinete da Secretária de Estado da Justiça, o Centro de Estudos Judiciários, a Comissão de Proteção às Vítimas de Crimes, a Comissão de Programas Especiais de Segurança, a Comissão de Apreciação e Controlo da Atividade dos Administradores de Insolvência e o Conselho Consultivo da Justiça.

24.000 trabalhadores, com uma taxa de habilitação superior em torno dos 20%, sendo cerca de 56% do total de trabalhadores do sexo feminino (Correia, 2015).

Em 2008, e em observância à Lei n.º 66-B/2007, de 28 de dezembro (Portugal, 2007), o Ministério da Justiça, como entidade enquadrada na administração pública portuguesa, instituiu a nova versão do SIADAP (processo de implementação ainda hoje em curso) que, *grosso modo*, pode ser entendido como um instrumento de medição no mérito do desempenho de serviços, dirigentes e trabalhadores.

Aos 10 serviços da administração direta e indireta do Estado anteriormente mencionados, é aplicável, por princípio, o SIADAP⁶. Fora do âmbito de aplicação do SIADAP ficam os tribunais (enquanto instituições) e os magistrados (judiciais e do Ministério Público⁷). No caso dos magistrados judiciais e do Ministério Público, a aferição do mérito no desempenho de funções encontra-se vertida, respetivamente nos Estatuto dos Magistrados Judiciais (Portugal, 1985) e Estatuto do Ministério Público (Portugal, 1986) (e subseqüentes alterações a estes estatutos), estando o Ministério da Justiça arredado desse procedimento.

No que respeita ao que tem sido a prática da aplicação do SIADAP aos serviços, isto é, à medição do mérito no desempenho no setor da justiça, destacam-se três importantes desafios a ser ultrapassados. São eles a medição de metas no cumprimento de objetivos temporais, a imposição de limites à superação de objetivos e a distribuição dos pesos pelos três vetores fundadores do SIADAP (eficácia, eficiência e qualidade). Estes desafios são particularmente interessantes por serem transversais a muitos, senão mesmo à maioria, dos demais setores de atividade.

O primeiro desafio, associado aos objetivos temporais, isto é, em que a meta consiste no cumprimento de uma tarefa bem delimitada até uma determinada data predefinida, pode ser bem ilustrado pelos objetivos negociados entre tutela e serviços para cumprimento até ao final de cada ano. Imagine-se, a título de exemplo, um serviço que tem de produzir um determinado relatório até à semana 52 de um determinado ano. Uma vez que, formalmente, passada a semana

⁶ Factualmente, o SIADAP é aplicado por nove desses serviços. A Polícia Judiciária, depois da aplicação do SIADAP em 2008, entendeu que o mesmo não lhe é aplicável (com base no disposto artigo 20.º da Lei n.º 49/2008, de 27 de agosto (Portugal, 2008a) e no artigo 36.º da Lei n.º 53/2008, de 29 de agosto (Portugal, 2008b)) e suspendeu a sua aplicação, situação que se mantém até hoje.

⁷ Em Portugal, segundo as estatísticas oficiais (disponíveis em www.dgpi.mj.pt), existem cerca de 1.800 magistrados judiciais e cerca de 1.400 magistrados do Ministério Público.

52 se dá início a um novo ano civil, em caso de desvio, mesmo que marginal, esse serviço arrisca o incumprimento total do respectivo objetivo.

O segundo desafio, resultante da observação de alguns desvios consideráveis (no sentido das superações), nos primeiros anos de aplicação da nova versão do SIADAP no Ministério da Justiça, fruto em grande medida de uma definição de metas demasiadamente conservadoras⁸, pode ser percebido como um imperativo de definição de limites máximos de superação, por forma a garantir um aperfeiçoamento continuado da forma como os valores constantes das metas de cada objetivo são estimadas.

As soluções para estes dois desafios surgem com a introdução, em 2010 e 2011, de duas inovações apresentadas pelo Grupo de Trabalho do Conselho Coordenador da Avaliação de Serviços – Rede GPEARL.

A primeira dessas soluções consistiu na introdução da seguinte fórmula de cálculo para obtenção da taxa de realização de um objetivo:

Fórmula 1 – Fórmula de cálculo para obtenção da taxa de realização de um objetivo

$$\text{Taxa de realização (\%)} = 100 + |\text{Resultado} - \text{Meta}| \times \frac{25}{|\text{Valor crítico} - \text{Meta}|}$$

Fonte: adaptado de Ministério das Finanças (2010).

Esta fórmula introduz, no SIADAP, a ideia de que os objetivos podem apresentar não só metas pontuais mas também metas por intervalo, em que existem tolerâncias associadas⁹. Contribui também para a introdução da ideia de *benchmark*, um valor de excelência, convencionando-se que, à falta de melhor referência histórica nacional ou internacional, deverá corresponder a uma superação de 125%. É a esse *benchmark* que se dá o nome de valor crítico. As vantagens desta inovação são evidentes:

⁸ Uma vez que os Quadros de Avaliação e Responsabilização (QUAR) são publicados nos sítios de internet dos serviços, é possível encontrar, *on line* e para cada serviço, o histórico (de 2007 em diante) dos resultados obtidos, globalmente e, de forma discriminada, para cada objetivo.

⁹ A definição de objetivos passa, desta forma, a poder ser feita com base na definição de metas pontuais, em que quaisquer desvios positivos ou negativos corresponderão necessariamente a superações ou incumprimentos parciais dos objetivos ou, ao invés, a poder ser feita com base em intervalos de valores: quaisquer resultados obtidos que possam ser localizados no interior dos intervalos definidos correspondem a cumprimentos desses mesmos objetivos, não podendo, no entanto, ser considerados como superações ou incumprimentos – surge a ideia de uma margem de tolerância.

- Permite conferir maiores margens de segurança na definição de metas para objetivos;
- Permite mitigar a proliferação de taxas de realização desadequadas resultantes da tendência para a subavaliação de metas;
- Permite utilização versátil o suficiente para acomodar objetivos de incremento positivo e negativo;
- Permite às entidades avaliadoras uma interpretação mais clara dos desvios positivos;
- Permite uma mais fácil e imediata comparação entre unidades avaliadas.

Mais tarde, em 2011, em reunião de 8 de abril, o Grupo de Trabalho do Conselho Coordenador da Avaliação de Serviços fixa como máximo para a taxa de realização de indicadores e objetivos o valor de 135% (não alterando os 125% associados ao conceito de valor crítico).

Desta forma, os dois primeiros desafios à medição do mérito no desempenho ficam superados. Por um lado, é reforçada a ideia de que, provavelmente, a melhor forma de tratar um objetivo temporal cuja meta esteja próxima ao final do ano civil passa pela definição de uma meta por intervalo¹⁰. Por outro, pelo que foi acima exposto, as superações excessivas passaram a estar formalmente impedidas.

Por fim, é importante dar destaque ao terceiro desafio, que consiste na distribuição dos pesos pelos três vetores fundadores do SIADAP: eficácia, eficiência e qualidade. Não havendo qualquer regra balizadora, e mesmo tendo em conta que não é possível identificar um padrão, dada a falta de homogeneidade na definição dos pesos, de entidade avaliada para entidade avaliada, é possível notar um padrão de atuação macro. Por regra, e mesmo não sendo possível quantificar a tendência, os serviços do Ministério da Justiça atribuem maiores pesos aos vetores da eficácia e da eficiência¹¹ do que ao vetor da qualidade. Este é um desafio ao qual, volvida quase uma década de aplicação do SIADAP nesta sua nova configuração, ainda não foi possível

¹⁰ Por exemplo, a um determinado objetivo temporal, em vez da atribuição de uma meta pontual correspondente à semana 52, poderá ser atribuída uma meta correspondente à semana 50, colocando-se a tolerância em 2 semanas. Neste exemplo, qualquer resultado situado entre a semana 48 e a semana 52 corresponde ao cumprimento do respetivo objetivo.

¹¹ A variabilidade entre a distribuição dos pesos destes dois vetores é tal que não permite o isolamento de nenhuma tendência adicional.

dar resposta adequada. Os motivos pelos quais este fenómeno é detetável nos QUARs dos serviços podem ser variados e constituem certamente um motivo de interesse para a comunidade científica, deixando-se aqui o mote para o aprofundamento desta questão. Contudo, e em jeito de conjectura, é possível avançar que os motivos para este fenómeno talvez possam encontrar raiz numa administração pública e, em particular, numa administração da justiça ainda demasiadamente presa aos preceitos weberianos da competência técnica neutral, focada na eficácia e na eficiência, e ainda pouco sensível a abordagens mais recentes como a nova gestão pública ou a *governance*, mas orientadas para a qualidade, isto é, para uma obtenção de resultados que não descarta o foco no cidadão (cliente) (Correia, 2016).

REFERÊNCIAS

- Altbach, P. (2008). The complex roles of universities in the period of globalization. In GUNI (ed.) *Higher Education in the World: New Challenges and Emerging Roles for Human and Social Development* (pp. 5-14). London: Palgrave Macmillan.
- Bilhim, J. (2012). O Mérito nos Processos de Seleção da Alta Direção da Administração Pública Portuguesa: Mito ou Realidade? *Seqüência: Estudos Jurídicos e Políticos*, Vol. 33 (65): 57-78.
- Bilhim, J. (2013). Papel dos Gestores na Mudança Cultural da Administração Central do Estado: O Caso da Meritocracia. *Passagens Revista Internacional de História Política e Cultura Jurídica*, Vol. 5 (2): 205-227.
- Bilhim, J.; Correia, P. (2016). Diferenças nas Percepções dos Valores Organizacionais dos Candidatos a Cargos de Direção Superior na Administração Central do Estado. *Sociologia*, n. 31: 81-105.
- Bilhim, J.; Garcia, J. (2013). La Racionalidad Política y Técnica en la Selección de la Alta Administración Pública y Gestión Pública Portuguesa. *El Cronista*, n. 38: 36-52.
- Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.
- Braskamp, L.; Ory, J. (1994). *Assessing faculty work*. San Francisco, CA: Jossey- Bass.
- Castelar, I.; Veloso, A.; Ferreira, R.; Soares, I. (2010). Uma Análise dos Determinantes de Desempenho em Concurso Público. *Economia Aplicada*, Vol. 14 (1): 81-98.
- Cave, M.; Hanney, S.; Kogan, M.; Trevett, G. (1988). *The Use of Performance Indicators in Higher Education*. London: Jessica Kingsley.
- Coolican, H. (2014). *Research Methods and Statistics in Psychology. 6th Edition*. Hove: Psychology Press.
- Correia, P. (2011). Sobre o SIADAP, a Teoria Institucional e o Ministério da Justiça em Portugal. *Scientia Iuridica*, Vol. 60 (325): 101-130.
- Correia, P. (2012). O Impacto do Sistema Integrado de Gestão e Avaliação do Desempenho da Administração Pública (SIADAP) na Satisfação dos Colaboradores – O Caso dos Serviços do Ministério da Justiça em Portugal. Tese de Doutoramento, Universidade Técnica de Lisboa.

- Correia, P. (2015). Análise de Dados sobre a Problemática do Género na Gestão de Recursos Humanos dos Serviços do Ministério da Justiça em Portugal: Um Ministério no Feminino? *Scientia Iuridica*, Vol. 64 (337): 97-110.
- Correia, P. (2016). Ministério Público n.0 – Sobre como Institucionalizar uma Mentalidade Orientada para a Qualidade. In: *Qualidade na Justiça, Qualidade da Democracia*. Lisboa: Sindicato dos Magistrados do Ministério Público.
- Correia, P.; Moreira, M. (2016). Ministério da Justiça Versão 2.0: Sobre a Sociedade de Informação, os New Media e o Ministério da Justiça em Portugal. *Lex Humana*, Vol. 8 (2): 97-119.
- Dolton, P.; McIntosh, S.; Chevalier, A. (2003). *Teacher pay and performance* (Vol. 19). Inst of Education.
- Embiruçu, M.; Fontes, C.; Almeida, L. (2010). An indicator for the evaluation of the teaching performance in teaching institutions. *Ensaio: Avaliação e Políticas Públicas em Educação*, Vol. 18 (69): 795-820.
- Escrigas, C., Granados, J., Tandoon, R., e Hall, B. L. (2014). Editor's Introduction. Knowledge, Engagement and Higher Education: Contributing to Social Change. *Report: Higher Education in the World 2014* (pp. xxxi-xxxix). Global University Network for Innovation – GUNi.
- Marsh, H.; Roche, L. (1997). Making students evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist*, Vol. 52 (11): 1187-1197.
- Marston, W. (1928). *Emotions of Normal People*. London: Kegan Paul, Trench, Trubner & Co. Ltd.
- Ministério das Finanças (2010). *Avaliação dos Serviços – Linhas de Orientação Gerais*. Lisboa: Grupo de Trabalho do Conselho Coordenador da Avaliação de Serviços – Rede GPEARl.
- Portugal (1985). Lei n.º 21/85, de 30 de julho. *Diário da República*, I Série, n. 173: 1-23.
- Portugal (1986). Lei n.º 47/86, de 15 de outubro. *Diário da República*, I Série, n. 238: 2999-3124.
- Portugal (2007). Lei n.º 66-B/2007, de 28 de dezembro. *Diário da República*, 1.ª série, n. 250: 2-21.
- Portugal (2008a). Lei n.º 49/2008, de 27 de agosto. *Diário da República*, 1.ª série, n. 165: 6038-6042.
- Portugal (2008b). Lei n.º 53/2008, de 29 de agosto. *Diário da República*, 1.ª série, n. 167: 6135-6141.

- Portugal (2008c). Lei n.º 64-A/2008, de 31 de dezembro. *Diário da República*, 1.ª série, n. 252: 2-389.
- Portugal (2009). Lei n.º 205/2009, de 31 de Agosto. *Diário da República*, 1.ª série, n. 168: 5729-5757.
- Portugal (2010). Lei n.º 55-A/2010, de 31 de dezembro. *Diário da República*, 1.ª série, n. 253: 2-324.
- Portugal (2011). Lei n.º 64/2011, de 22 de dezembro. *Diário da República*, 1.ª série, n. 252: 42-240.
- Portugal (2012). Lei n.º 66-B/2012, de 31 de dezembro. *Diário da República*, 1.ª série, n. 250: 2-21.
- Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education*, Vol. 16 (2): 129-150.
- Ramos, M.; Moraes, R. (2000). Avaliação do desempenho de professores numa perspectiva qualitativa: contribuições para o desenvolvimento profissional de professores universitários. *Revista Iberoamericana de Educación*, n. 108: 1-18.
- Ramos Pinto, R. (2012). *Introdução à Análise de Dados. 2ª Edição*. Lisboa: Edições Sílabo.
- Wimmer, R.D. e Dominick, J.R. (2011). *Mass Media Research: An Introduction*. 9th Edition. Boston: Cengage Learning.

Universidade Católica de Petrópolis
Centro de Teologia e Humanidades
Rua Benjamin Constant, 213 – Centro – Petrópolis
Tel: (24) 2244-4000
lexhumana@ucp.br
<http://seer.ucp.br/seer/index.php?journal=LexHumana>



RAMOS PINTO, Ricardo; CORREIA, Pedro Miguel Alves Ribeiro. Mérito, desempenho e resultados: os desafios da medição. *Lex Humana*, v. 10, n. 2, p. 87-116, dez. 2018. ISSN 2175-0947. Disponível em: <<http://seer.ucp.br/seer/index.php/LexHumana/article/view/1587>>
