

UNIVERSIDADE DE LISBOA

INSTITUTO DE EDUCAÇÃO



LISBOA

UNIVERSIDADE
DE LISBOA

PARA UMA AVALIAÇÃO DO PROJETO TESTES INTERMÉDIOS

Maria Antonieta da Costa Mourão Lima Ferreira

Dissertação

MESTRADO EM CIÊNCIAS DA EDUCAÇÃO

Área de especialização em Avaliação em Educação

2015

UNIVERSIDADE DE LISBOA

INSTITUTO DE EDUCAÇÃO



PARA UMA AVALIAÇÃO DO PROJETO TESTES INTERMÉDIOS

Maria Antonieta da Costa Mourão Lima Ferreira

Dissertação de Mestrado orientada pelo Professor Doutor Domingos Fernandes

MESTRADO EM CIÊNCIAS DA EDUCAÇÃO

Área de especialização em Avaliação em Educação

2015

O som do silêncio

Devagar, como se tivesse todo o tempo do dia,
descasco a laranja que o sol me pôs pela frente. É
o tempo do silêncio, digo, e ouço as palavras
que saem de dentro dele, e me dizem que
o poema é feito de muitos silêncios,
colados como os gomos da laranja que
descasco. E quando levanto o fruto à altura
dos olhos, e o ponho contra o céu, ouço
os versos soltos de todos os silêncios
entrarem no poema, como se os versos
fossem como os gomos que tirei de dentro
da laranja, deixando-a pronta para o poema
que nasce quando o silêncio sai de dentro dela.

Nuno Júdice, *Guia de Conceitos Básicos*

AGRADECIMENTOS

Ao Professor Doutor Domingos Fernandes, o meu agradecimento pela orientação deste trabalho. O seu discurso e a sua prática de exigência aliados ao seu profundo saber sobre avaliação constituíram o incentivo certo para a prossecução do projeto.

Este trabalho não teria sido possível se eu não acreditasse na máxima de que nos esforçamos para fazer bem, antes de tudo, porque sim. Esta lição aprendi-a de meus Pais, a quem agradeço.

Também em nome desse valor, quero aqui registar o meu agradecimento pela honra que tive em partilhar a missão de serviço público com muitos que ao longo dos últimos anos comigo formaram equipa, trazendo diariamente o saber, a persistência, o rigor e a responsabilidade que o trabalho na esfera pública exige.

Daqueles destaco a Ana Costa e o Paulo Tapadas por, no início, me terem ensinado, cada um de forma diferente, a necessidade de se saber antes de formar convicção, de duvidar para mais aprender e de ter a inteligência e a honestidade do rigor.

Este trabalho não teria sido este sem a Vanda Lourenço. Pela generosidade na partilha de saber, pela perseverança no percurso, pelo espírito aguçado e inteligente, que se revela na excelência do profissionalismo e da amizade.

Aos amigos que me dão energia e muito especialmente às minhas amigas Anabela Carvalho, Anabela Neves, Filipa Barreto, Filomena Gamelas e Rita Melancia por constituírem a *amostra certa* do que é ser Professor. E me fazerem acreditar que devemos, antes de mais, trabalhar para os que fazem bem.

Por fim, aos meus filhos Maria e Nuno, que sempre souberam acrescentar, e bem, outros desafios a esta fase em que o trabalho decorreu e que me dão a imensa satisfação de se entregarem ao que fazem e de não se esquecerem de ser boas pessoas. Ao Rui, por estabelecer critérios elevados no seu escrutínio, pela determinação e inteligência. Porque exigimos sempre mais de quem mais gostamos.

RESUMO

Este estudo teve como objetivo a descrição e caracterização da participação das escolas no Projeto Testes Intermédios, no período 2010-2013, sistematizando informação e interpretando as taxas de adesão e a composição anual do elenco de testes intermédios. Pretendeu-se caracterizar o universo de escolas que aderiram ao Projeto e a natureza dessa adesão.

Desenvolveu-se uma investigação de cariz quantitativo, centrada na análise do comportamento das escolas perante a oferta de testes intermédios dirigida ao 3.º ciclo do ensino básico e ao ensino secundário, tendo como fonte primária os dados referentes à inscrição das escolas no período acima indicado. A análise partiu do conjunto inicial de variáveis de caracterização das escolas, a que acresceram as variáveis relativas aos anos de participação no Projeto e ao elenco de testes selecionado, como forma de identificar padrões de adesão.

Enquanto contribuição para a avaliação do Projeto Testes Intermédios, o estudo constitui-se como uma avaliação de uma intervenção de política pública educativa.

Os resultados evidenciaram uma participação muito expressiva, de dimensão nacional, consistente ao longo do triénio e com padrões de participação que revelarão a vontade de tirar partido das potencialidades de instrumentos de avaliação externa como moderação das avaliações internas. Esta expressão de vontade não se limitou a áreas disciplinares sujeitas a provas de avaliação externa, abrangeu outras áreas, nomeadamente aquelas em que, exceção feita ao período 2010-2013, se regista ausência de qualquer processo de aferição de âmbito nacional, através de avaliação externa.

Estas conclusões, consideradas por si só, fundamentariam uma continuidade do Projeto.

Palavras-chave: Projeto Testes Intermédios; Avaliação de Programas; Avaliação Externa; Avaliação Interna; Política Pública Educativa.

ABSTRACT

The main objective of this study was to describe and characterize schools' participation in Intermediate Tests Project from 2010 to 2013, synthesizing data and interpreting the participation rates and the annual composition of the available set of intermediate tests. The aim was to characterize the universe of schools that entered the project, as well as the nature of such participation.

The research followed a quantitative approach, centred on the analysis of the schools' behaviour against the range of intermediate tests available to lower (grades 7 to 9) and upper secondary (grades 10 to 12) levels of schooling. The main data source was the enrolment of schools in the Project, in the period mentioned above. The starting point was the initial set of school characterization variables, to which variables related to the years of Project participation were added as well as information on the *corpus* of selected tests, with the objective of identifying patterns of adherence.

While a contribution to the evaluation of Intermediate Tests Project, the study undertakes an evaluation of a public educational policy.

The results showed a very extensive participation, at national level, consistent throughout the three-year period and with adherence patterns that suggest schools take advantage of the potential benefits of external assessment as a mean of internal assessment moderation. This willpower was not limited to subject areas where the external assessment is mandatory; it was extended to areas in which there has been a total absence of any external assessment process, except from 2010 to 2013.

The above conclusions would, *per se*, support the continuation of the Intermediate Tests Project.

KEYWORDS: Intermediate Test Project; Program Evaluation; External Evaluation; Internal Evaluation; Public Educational Policy.

ÍNDICE

Agradecimentos	i
Resumo	ii
Abstract.....	iii
Índice	iv
Índice de Gráficos.....	vi
Índice de Quadros.....	vii
Lista de abreviaturas e siglas	viii
Capítulo 1 – Enquadramento e Questões de Investigação.....	1
Introdução.....	1
Relevância e Pertinência da Investigação.....	2
Questões e Matriz da Investigação	5
Capítulo 2 – Objeto de estudo: o Projeto Testes Intermédios	8
Definição e Caracterização	8
Capítulo 3 – Revisão da literatura	20
Avaliação em Educação.....	20
Avaliação no domínio das aprendizagens: Avaliação interna.....	27
Avaliação no domínio das aprendizagens: Avaliação externa	35
Capítulo 4 – Metodologia.....	42
Opções Metodológicas e Procedimentos Principais.....	42
Capítulo 5 – Apresentação e Discussão dos Resultados.....	45
A Base de Dados do Projeto Testes Intermédios (Triénio 2010-2013).....	45
O Universo de Escolas Participantes no Projeto Testes Intermédios	46
A Adesão das Escolas ao Projeto	56
O Universo de Escolas de Tipologia de Participação TI EBS.....	61
Caracterização das escolas TI EBS	62
Padrões de Adesão.....	67
Capítulo 6 – Conclusão e Reflexões Finais	73
Nota Prévia	73
Síntese dos Principais Resultados	74
Reflexões a Partir da Síntese de Resultados.....	77
Nota Final	86
Glossário.....	87
Bibliografia.....	88

Anexos.....	93
Anexo A.....	94
Anexo B.....	98
Anexo C.....	101
Anexo D.....	102
Anexo E.....	104

ÍNDICE DE GRÁFICOS

Gráfico 1 – Oferta de número de disciplinas e de TI, por ano (2005-2015)	11
Gráfico 2 – Oferta de número de disciplinas e de TI, por ano (2010-2013)	12
Gráfico 3 – Número de disciplinas no Projeto TI com avaliação externa e sem avaliação externa, por ano (2010-2013)	12
Gráfico 4 – Número de provas disponibilizadas e média de inscrições por escola no triénio (2010-2013)	47
Gráfico 5 – Escolas participantes no Projeto TI, por ano e natureza jurídica (em %)	50
Gráfico 6 – Ordenação da média de inscrições no Projeto TI por NUTS III (2010-2013) .	55
Gráfico 7 – Testes realizados nas Escolas TI EBS no triénio 2010-2013 (em %).	64
Gráfico 8 – Disciplinas cobertas nas Escolas TI EBS no triénio 2010-2013 (em %).	65
Gráfico 9 – Relação entre a quantidade de TI e a variedade de TI no triénio (2010-2013)	67

ÍNDICE DE QUADROS

Quadro 1 – Matriz de investigação.....	7
Quadro 2 – Elenco de TI disponibilizados (disciplina e ano de escolaridade), por ciclo de escolaridade (3.º CEB e ES), por ano (2010-2013).....	14
Quadro 3 – Número de anos de participação das escolas no Projeto TI no triénio (2010-2013).....	46
Quadro 4 – Escolas participantes por Tipologia de Participação no triénio (2010-2013)...	49
Quadro 5 – Adesão média das escolas participantes no Projeto TI, por ano e natureza jurídica.....	50
Quadro 6 – Distribuição de escolas que realizaram Provas Finais do 9.º ano e Exames Nacionais do 12.º ano no triénio 2010-2013, por natureza jurídica.....	52
Quadro 7 – Média de inscrições por NUTS III no triénio 2010-2013.....	53
Quadro 8 – Adesão das escolas ao Projeto TI, por ano e ciclo de escolaridade.....	57
Quadro 9 – Número ordenado de inscrições, por Teste Intermédio e ciclo de escolaridade (2010/2011).....	59
Quadro 10 – Número ordenado de inscrições, por Teste Intermédio e ciclo de escolaridade (2011/2012).....	60
Quadro 11 – Número ordenado de inscrições por Teste Intermédio e ciclo de escolaridade (2012/2013).....	60
Quadro 12 – Média de inscrições das Escolas TI EBS por natureza jurídica.....	62
Quadro 13 – Média de inscrições das Escolas TI EBS que participaram nos três anos do Projeto TI, por NUTS III.....	63
Quadro 14 – Projeto TI – Padrões de seleção.....	71

LISTA DE ABREVIATURAS E SIGLAS

1.º CEB	1.º Ciclo do Ensino Básico
3.º CEB	3.º Ciclo do Ensino Básico
BD	Base de dados
BD TI	Bases de dados do Projeto Testes Intermédios
EPE	Escolas Portuguesas no Estrangeiro
ES	Ensino Secundário
Ex	Exame Final Nacional
IAVE, I.P.	Instituto de Avaliação Educativa, I.P. (Serviço que sucedeu, após setembro de 2013, ao Gabinete de Avaliação Educativa, GAVE)
OCDE	Organização para a Cooperação e Desenvolvimento Económico
PF	Prova Final Nacional
Projeto TI	Projeto Testes Intermédios ¹
TI	Testes intermédios
TI – Testes intermédios por disciplina² e ano de escolaridade	
BG_10	Biologia e Geologia – 10.º ano
BG_11_t1	Biologia e Geologia – 11.º ano – teste 1
BG_11_t2	Biologia e Geologia – 10.º ano – teste 2
CFQ_9/FQ_9	Ciências Físico-Químicas/Físico-Química – 9.º ano
CN_9	Ciências Naturais – 9.º ano
Filos_10	Filosofia – 10.º ano
Filos_11	Filosofia – 11.º ano
FQ_A_10	Física e Química A – 10.º ano
FQ_A_11_t1	Física e Química A – 11.º ano – teste 1
FQ_A_11_t2	Física e Química A – 11.º ano – teste 2
Geo_9	Geografia – 9.º ano
Hist_9	História – 9.º ano
Ing_9	Inglês – 9.º ano
LP_9/Port_9	Língua Portuguesa/Português – 9.º ano
Mat_A_10	Matemática A – 10.º ano
Mat_A_11_t1	Matemática A – 11.º ano – teste 1
Mat_A_11_t2	Matemática A – 11.º ano – teste 2
Mat_A_12_t1	Matemática A – 12.º ano – teste 1
Mat_A_12_t2	Matemática A – 12.º ano – teste 2

¹ No presente trabalho, a referência ao Projeto exclui os testes intermédios que, a partir de 2010/2011, se dirigem ao 1.º ciclo do ensino básico, 2.º ano.

² Na sequência do disposto no Decreto-Lei n.º 139/2012, de 5 de julho, a designação das disciplinas de Ciências Físico-Químicas e de Língua Portuguesa passa a ser Físico-Química e Português, respetivamente.

Mat A_12_t3	Matemática A – 12.º ano – teste 3
Mat B_10	Matemática B – 10.º ano
Mat B_11	Matemática B – 11.º ano
Mat_8	Matemática – 8.º ano
Mat_9_t1	Matemática – 9.º ano – teste 1
Mat_9_t2	Matemática – 9.º ano – teste 2
Port_12	Português – 12.º ano

CAPÍTULO 1 – ENQUADRAMENTO E QUESTÕES DE INVESTIGAÇÃO

INTRODUÇÃO

O presente estudo resulta da percepção de que o Projeto Testes Intermédios (TI) assumia um lugar significativo no conjunto de intervenções que, na área da avaliação das aprendizagens, o Ministério da Educação e Ciência desenvolveu nos últimos dez anos.

A percepção sobre a sua dimensão e importância conduziu à necessidade de compreender a forma como, no seu desenvolvimento, o Projeto prosseguiu o quadro de finalidades que o definiu.

A evidente complexidade da avaliação de um projeto de âmbito nacional, tendo como base uma parceria entre um serviço central e as escolas da rede escolar portuguesa, relançado a cada ano letivo e dirigido à avaliação das aprendizagens, exigiu, no âmbito dos objetivos de um estudo desta natureza, que fossem tomadas opções que permitissem uma abordagem rigorosa e exequível.

As escolhas realizadas relativamente à natureza dos dados a analisar decorreram da opção de, perante a multiplicidade de perspetivas de análise possíveis, se ter querido *começar pelo princípio*, isto é, começar por descrever e caracterizar o fenómeno da adesão das escolas ao Projeto, porque havia evidências claras da sua popularidade. A descrição e caracterização da adesão mereciam atenção especial uma vez que, neste Projeto, as escolas participavam de forma voluntária e o padrão da sua participação era igualmente construído em função das suas decisões.

Dar significado às taxas de inscrição e à composição do elenco de testes intermédios a realizar, ao longo da série temporal em estudo, constituiu-se como o objetivo central, sabendo desde o início que o estudo se configuraria apenas como uma contribuição, embora estrutural, para a avaliação do Projeto.

O texto que, de seguida, se apresenta, está estruturado em seis grandes partes, correspondentes a seis capítulos. No Capítulo 1, enquadra-se o estudo realizado, suportando a sua relevância e pertinência e definem-se os objetivos e questões da investigação. No Capítulo 2, caracteriza-se o Projeto Testes Intermédios enquanto objeto de estudo. O Capítulo 3 apresenta a revisão da literatura. O Capítulo 4 é dedicado à metodologia, explicitando as principais opções e procedimentos seguidos. O Capítulo 5 faz a apresentação e discussão dos resultados, a que se seguem as conclusões e reflexões finais, no Capítulo 6.

RELEVÂNCIA E PERTINÊNCIA DA INVESTIGAÇÃO

É unanimemente aceite que se desenvolveram alterações significativas no sistema educativo português durante as últimas décadas. São muitos e diversos os indicadores que revelam profundas transformações neste período. A caracterização global daquelas alterações, que não cabe, naturalmente, neste projeto de estudo, teria de enquadrar as evoluções em várias áreas do complexo sistema nacional de ensino, que envolvem um conjunto de atores, instituições e recursos que concorrem para que o Estado assegure o serviço público da Educação através do desenvolvimento de política educativa e da administração dos vários serviços.

Esta já longa etapa de transformação global integra, obviamente, intervenções na área da avaliação, com especial incidência nos últimos catorze anos³. Neste período, há mudanças assinaláveis, fundamentalmente nos domínios da avaliação do desempenho de alunos, da avaliação de professores e, também, da avaliação do sistema (escolas e políticas educativas).

O relatório *OECD Reviews of Evaluation and Assessment in Education: Portugal* (Santiago, P.; Donaldson, G.; Looney, A., & Nusche, D., 2012)⁴ caracteriza a década 2002-2012 como

³ Serão catorze anos se considerarmos o primeiro ciclo PISA (PISA 2000) ou a aplicação de provas de aferição pela primeira vez em 2000.

⁴ Em 2013, a OCDE publicou o relatório *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, em que colige e compara a experiência de 28 países da OCDE no que respeita a

a que, em Portugal, corresponde à das *fundações* de um “framework for evaluation and assessment”⁵ (p. 9). Destacando sistematicamente o significativo desenvolvimento português neste domínio, o Relatório afirma que há um conjunto de iniciativas que “clearly communicate that evaluation and assessment are priorities in the school system and reveal a coherent and comprehensive agenda to develop an evaluation culture among school agents” (p. 9). Simultaneamente, a equipa responsável pelo estudo identifica um conjunto de desafios a que esta agenda tem de responder, já que é necessário o reforço de alguns dos componentes do referido “framework for evaluation and assessment”. Afirma-se que é necessário que se assegure articulação dentro do quadro de referências da avaliação, para que se atinja consistência e complementaridade na intervenção e para que as práticas avaliativas sejam orientadas para a melhoria⁶.

O Relatório OCDE (Santiago *et al.*, 2012) especifica, no contexto do estudo que lhe está na base, a que se refere e o que compreende o termo avaliação (no original, *assessment*): por avaliação entendem-se os julgamentos sobre progresso e cumprimento de objetivos de aprendizagem de alunos (tomados individualmente), abrangendo avaliações de sala de aula (avaliação interna), avaliações de larga-escala, testes externos e exames. Isto é, instrumentos de natureza diversa e com propósitos distintos.

No âmbito dos desafios que o Relatório globalmente identifica no domínio da avaliação (*assessment*), são elencadas prioridades, sendo que as relativas a *The evaluation and assessment framework* e *Student assessment* constituem matéria intrinsecamente relacionada com o Projeto TI, objeto central do estudo que agora se apresenta.

Durante o desenvolvimento deste trabalho, e nomeadamente no período coincidente com a análise de dados, o Projeto TI foi objeto de uma reconfiguração profunda, que culminou, em

políticas nacionais de avaliação educativa (ensino primário e secundário). O Relatório de 2012 relativo a Portugal é o relatório prévio que, no caso português, informa o relatório final (2013).

⁵ O Relatório usa os termos *assessment*, *appraisal* e *evaluation* para se referir, respetivamente, a avaliação de alunos, de professores e de sistema.

⁶ Perspetiva profundamente comprometida com a defesa de uma agenda em que a avaliação serve a aprendizagem, na medida em que o processo educativo se orienta para a melhoria da qualidade.

setembro de 2014, com o anúncio da sua descontinuidade, nos termos que serão oportunamente explicitados.

A necessidade de se conhecerem de forma sistemática, para se avaliarem, as intervenções de políticas públicas é um desiderato que se mantém inalterado na sua essência perante a interrupção de uma determinada intervenção, programa ou projeto. Mais, se o quadro de intervenção for o das opções sustentadas em evidências partilhadas, seria sempre desejável que a avaliação precedesse a decisão. Como se verá adiante, os fundamentos do juízo que terá levado à decisão de descontinuidade do Projeto TI não são evidentes.

O projeto de investigação que deu origem a este trabalho ganhou, assim, maior relevância e pertinência. Há dados que não se conheciam e que se pretendeu que ficassem sistematizados, como base para outros estudos, com a finalidade de contribuir para o que efetivamente se sabe sobre uma intervenção dirigida à melhoria da aprendizagem dos alunos, da responsabilidade de um organismo que cumpre, desde 2005/2006 (ano letivo do lançamento do Projeto), uma missão superiormente definida pelo Ministério competente, um projeto que envolveu mais de um milhar de escolas e centenas de milhares de alunos⁷ e que constituiu certamente um investimento significativo de recursos humanos e financeiros.

Embora se reconheça, por tudo o que já se explicitou, fundamentação bastante para o presente trabalho, a coincidência no tempo da sua concretização com a descontinuidade do Projeto TI, no quadro conceptual em que havia sido criado, trouxe a oportunidade de poder contribuir para avaliar a pertinência do momento da suspensão.

⁷ Veja-se, a título de exemplo, que no ano letivo 2011/2012, ano em que a oferta cobriu 18 disciplinas, os alunos envolvidos realizaram 888 844 testes (538 336 no 3.º CEB e 350 508 no ES), dados disponíveis nos documentos denominados *Informações – Resultados nacionais e regionais 2011/2012* (acessíveis a diretores, gestores de projeto e professores, através da Extranet do IAVE, I.P.).

QUESTÕES E MATRIZ DA INVESTIGAÇÃO

Entendendo-se este Projeto como concretização de um programa de política pública de educação, reafirma-se a instância da sua avaliação, visando-se julgar a forma como tem conseguido responder às finalidades a que se propõe e satisfazer as necessidades dos potenciais beneficiários (alunos, professores e escolas). No caso concreto do presente estudo pretendeu-se traçar uma linha de investigação direcionada para a avaliação de uma das vertentes distintivas do Projeto TI: a opcionalidade da adesão das escolas.

Efetivamente, a facultatividade da inscrição das escolas, bem como a possibilidade de cada uma constituir, a cada ano, o elenco de testes intermédios que pretendia aplicar são características a que o grupo de trabalho da OCDE responsável pelo Relatório acima referenciado (Santiago *et al.*, 2012) não foi insensível. Na página 54 do Relatório, encontra-se uma referência a estes elementos de avaliação como “[t]he optional but very popular intermediate tests”. Porém, após explicitação da sua natureza e finalidades (com relevo legítimo para as de *suporte ao processo avaliativo*, de aferição de padrões nacionais de desempenho e de regulação, para os alunos, da sua aprendizagem), o texto do Relatório conclui: “The majority of schools use them as summative assessment instruments.” (p. 54). E a sua utilização (quase apenas) como instrumentos de avaliação sumativa configurará um aproveitamento insuficiente destes instrumentos. A observação aqui reproduzida não deixa de constituir um reforço adicional à necessidade de avaliação que o Projeto, por si, exige.

Centrado na primeira questão — o facto de a sua facultatividade conviver com a sua popularidade (evidente, por exemplo, nos números totais de inscrições) —, o estudo teve de se focar nos dados que constituem e caracterizam o universo em análise: as escolas que aderiram ao Projeto e o elenco de TI que decidiram aplicar em cada ano. Tendo por referência o período que corresponde aos anos letivos 2010/2011, 2011/2012 e 2012/2013,

o triénio objeto de análise⁸, a investigação foi orientada pelas questões que a seguir se formulam:

- Como se poderá caracterizar o universo de escolas que aderiram ao Projeto TI no triénio 2010-2013?
- Como se poderá caracterizar a natureza da adesão das escolas ao Projeto?
 - Qual a quantidade e variedade de TI escolhidos pelas escolas?
 - Que padrões de participação (traduzidos na constituição do elenco de TI) poderão ser identificados?

Face a este enquadramento, julgou-se, então, válido começar por descrever e caracterizar o universo de escolas participantes no Projeto, para, de seguida, focar o estudo na descrição da qualidade⁹ da adesão das escolas ao projeto TI, com vista a desenhar padrões de participação.

Em suma, e tendo a *adesão das escolas ao Projeto* como objeto de avaliação, pretendeu-se, através deste estudo *Descrever, analisar e interpretar a adesão das escolas ao Projeto TI*.

Nessa perspetiva de análise, assumidamente parcial e limitada perante as ambiciosas finalidades que o Projeto define, visou-se sistematizar e analisar informação que possa vir a contribuir para a determinação do valor do Projeto, tendo em conta “a) os objetivos do programa; b) as necessidades dos principais destinatários do programa; c) os objetivos da medida de política a que o programa possa estar associado; e d) as preferências daqueles que, de algum modo, possam ser afetados pela avaliação e/ou que tenham interesse nos seus resultados, os chamados *stakeholders*.” (Fernandes, 2011, p. 190).

⁸ Período temporal a que correspondem os dados de inscrição disponibilizados, no âmbito de protocolo estabelecido entre o IAVE, I.P. e o IE-UL (documento reproduzido no Anexo A).

⁹ A palavra *qualidade* está a ser utilizada na aceção de “propriedade ou característica de uma coisa”. No caso concreto, refere-se ao conjunto de propriedades que definem a adesão de uma escola ao Projeto, i.e., quantos testes e que testes selecionam para o seu elenco de participação. Trata-se, portanto, a aceção etimológica da palavra (*qualitas, -atis*), que não implica uma descrição valorativa da adesão.

Identificado o objeto de avaliação, definido o objetivo e formuladas as questões de investigação, desenhou-se uma matriz de investigação, como a que se apresenta no Quadro 1.

Quadro 1 – Matriz de investigação

Objeto	Dimensões
Projeto TI — Adesão das escolas	Elenco de TI a aplicar em cada ano Evolução da adesão das escolas Distribuição da adesão Padrões de adesão

CAPÍTULO 2 – OBJETO DE ESTUDO: O PROJETO TESTES INTERMÉDIOS

DEFINIÇÃO E CARACTERIZAÇÃO

Nesta secção, apresenta-se e caracteriza-se o Projeto TI, tendo por referência o triénio 2010-2013, não deixando de recuperar o histórico desde o seu lançamento, em 2005/2006.

Enquanto intervenção da responsabilidade do Gabinete de Avaliação Educacional (GAVE)¹⁰ – um serviço central do Ministério da Educação e Ciência (MEC), o Projeto TI, iniciado, como referido anteriormente, em 2005/2006, foi desenvolvido no âmbito da missão consignada àquele Gabinete (missão presentemente da responsabilidade do IAVE, I.P.¹¹), constituindo uma atividade integrada no programa central daquele serviço – “contribuir para a melhoria da qualidade das aprendizagens dos alunos e das práticas dos professores suportadas pela aplicação de instrumentos de avaliação educacional de elevada qualidade técnica e científica, assegurando, em paralelo, a recolha de informação relevante para o diagnóstico do sistema educativo nacional e para a tomada de decisões que concorram para incrementar a sua eficiência”.¹²

Em termos gerais, o Projeto TI consiste na elaboração de instrumentos de avaliação (os testes intermédios) e na sua disponibilização às escolas que, com carácter facultativo, aderem ao Projeto, aplicando os testes aos seus alunos, de acordo com um conjunto de regras e de procedimentos que, no momento da inscrição, se comprometem a fazer cumprir. Os TI são, assim, instrumentos de avaliação, concebidos centralmente, para aplicação pelas escolas em datas pré-definidas do calendário escolar, com realização simultânea nas escolas aderentes, com classificação local, a partir de critérios de classificação da responsabilidade, naturalmente, da entidade que elabora os instrumentos. O resultado da classificação dos

¹⁰ Serviço a que sucedeu, após setembro de 2013, nos termos do Decreto-Lei n.º 102/2013, de 25 de julho, o Instituto de Avaliação Educativa, I.P. (IAVE, I.P.).

¹¹ No domínio específico da missão a que nos reportamos, não se registam alterações significativas decorrentes da criação do IAVE, I.P..

¹² Excerto do texto identificado como “Missão”, disponível em <http://www.gave.min-edu.pt/np3/2.html> (consultado em 14 de junho de 2013 e em 18 de novembro de 2014).

testes é devolvido, por cada escola, ao IAVE, I.P., que, por sua vez, compila e trata os dados, e apresenta análise nacional e regional de cada um dos TI.

É importante, ainda, referir a natureza *sui generis* que os testes intermédios apresentam enquanto instrumentos de avaliação concebidos centralmente, com aplicação nacional síncrona, e com utilização local, contextualizada no âmbito do processo de avaliação interna de cada escola. Estamos perante um instrumento de carácter híbrido, com características que o aproximam, por um lado, das avaliações *high-stakes*, não na essência ou funções, naturalmente, mas pela identidade de alguns traços, e, por outro, das provas *low-stakes*, pelo facto de não terem implicações (pré-determinadas, obrigatórias) no percurso escolar do aluno e por serem fonte (à partida, privilegiada) de informação de natureza reguladora. A estas especificidades juntam-se outras tantas, que parecem consagrar a atipicidade dos instrumentos de avaliação disponibilizados pelo Projeto. Veja-se, por exemplo, o facto de a aplicação de critérios de classificação ser local, ser realizada pelos próprios professores dos alunos, sem acompanhamento central de aferição ou de concertação, o facto de caber à escola e, em última análise ao professor da turma, a decisão sobre a forma de ponderar a informação avaliativa produzida por cada TI, ou, ainda, o facto de a informação devolvida pelo IAVE, I.P., apresentar análise de âmbito nacional com leituras comparativas de resultados regionais (também interanuais)¹³.

Os documentos de referência do Projeto¹⁴, divulgados na página eletrónica do IAVE, I.P. e na denominada Extranet (plataforma de acesso dedicado às direções das escolas), enunciam as finalidades que se reproduzem: “[a)] permitir a cada professor aferir o desempenho dos seus alunos por referência a padrões de âmbito nacional, [b)] ajudar os alunos a uma melhor consciencialização da progressão da sua aprendizagem e, [c)] complementarmente (sublinhado nosso), contribuir, sempre que aplicável, para a sua progressiva familiarização com instrumentos de avaliação externa, processo a que estarão sujeitos no final dos ciclos do ensino básico, ou no ano terminal das disciplinas do ensino secundário.” (GAVE, 2012).

¹³ Veja-se, sobre este último aspeto, o exemplo de uma Informação-Resultados (Anexo B).

¹⁴ Todas as referências feitas nesta secção têm por base os documentos publicados pelo GAVE para o ano letivo de 2012/2013.

A explicitação destas finalidades do Projeto é acompanhada de um conjunto de outras considerações que pretenderão clarificar os objetivos principais da intervenção. Nos mesmos documentos, afirma-se:

é importante salientar o carácter eminentemente formativo dos instrumentos de avaliação a disponibilizar. De facto, a aferição dos desempenhos e a regulação da aprendizagem, a partir de uma reflexão sustentada pela análise do processo de resposta dos alunos e pelos resultados atingidos, são objetivos centrais deste projeto, que deverão presidir à opção de cada escola no momento de adesão ao projeto e na escolha dos testes a realizar. (GAVE, 2012).

A dimensão e abrangência que este Projeto atingiu nos últimos anos¹⁵, nomeadamente no triénio 2010-2013, em complemento da sua natureza e propósitos, são traços distintivos de uma intervenção que se configura como de política educativa, implementada ao abrigo da missão consignada ao GAVE e de que o recém-criado IAVE, I.P. é herdeiro.

No Gráfico 1, fica evidente a evolução da oferta TI nos dez anos de existência do Projeto. Para o presente estudo, considera-se a oferta TI dirigida ao 3.º CEB e ES, não se incluindo a oferta dirigida ao 2.º ano do 1.º CEB, introduzida a partir de 2010/2011 (com os TI de Matemática e Português).

A distribuição ao longo da década permite a constatação daquilo que se poderiam considerar cinco fases de evolução; (i) uma primeira fase, que coincide com o lançamento do Projeto, em que a oferta quantitativa de TI é de baixa expressão (4 testes em 2005/2006 e 3 testes no ano seguinte) e é acompanhada de uma concentração em apenas duas disciplinas – Matemática A e Matemática B¹⁶; (ii) uma segunda etapa, que abrange os anos de 2007/2008

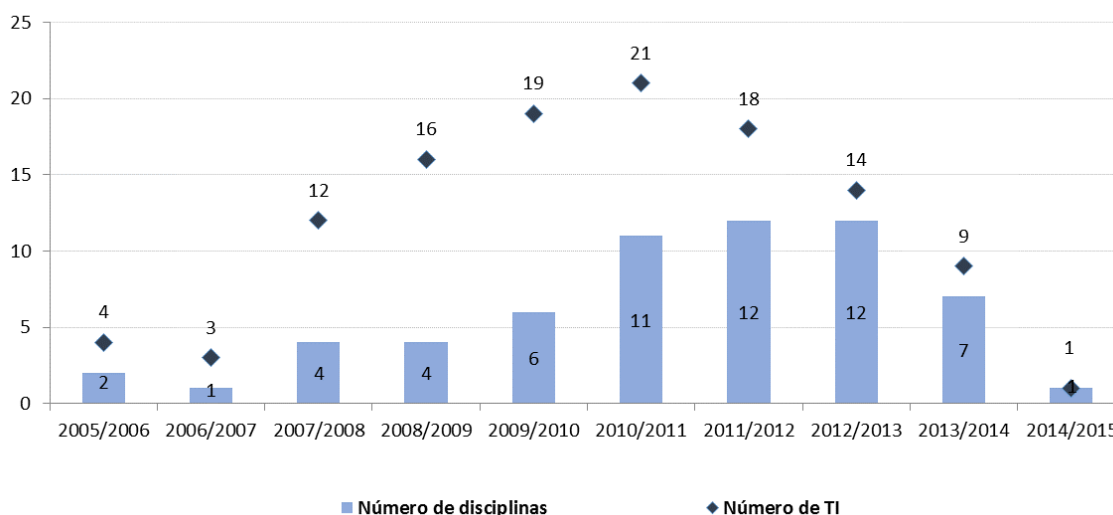
¹⁵ Se tomarmos como referência o Relatório Testes Intermédios 2012 (GAVE, 2013b) e as informações parciais de resultados (Informações-resultados: uma para cada TI aplicado a nível nacional), há um conjunto de dados que importará reproduzir pelo facto de estes serem indicadores da dimensão e da abrangência do Projeto. Em 2011/2012, foram disponibilizados 21 testes intermédios. Na disciplina de Matemática (9.º ano), o teste foi realizado por 91 051 alunos, na de Português, por 87 763 alunos (valores mais elevados de participação) e, na disciplina de História, que registou a menor taxa de participação, o teste foi realizado por 47 507 alunos. No ensino secundário, o teste da disciplina de Português (12.º ano) foi realizado por 48 097 alunos (o valor mais elevado), na de Matemática A (12.º ano – teste do 3.º período), por 31 584 alunos, sendo que, na disciplina de Biologia e Geologia (11.º ano), contou com 28 549 respondentes (o valor mais baixo).

Refira-se que aquele mesmo Relatório apresenta dados relativos à georreferenciação de resultados que permitem afirmar que todas as NUTS III que constituem o território nacional apresentavam escolas/agrupamentos de escolas inscritos no Projeto.

¹⁶ Consulte-se o Anexo C para identificação do elenco anual de TI para os anos 2005 a 2015.

a 2009/2010, em que há um aumento significativo do número de TI oferecidos, havendo um crescimento também muito significativo da oferta disciplinar de TI; (iii) uma terceira fase, que coincide com o triénio 2010-2013, em que a redução do número de testes é significativa (depois de atingido um pico em 2010/2011, a evolução é equivalente, mas com sentido contrário, à inscrita no triénio 2007-2010) e é acompanhada de uma clara diversificação da oferta disciplinar (11 disciplinas em 2010/2011 e 12 nos dois anos seguintes); (iv) um quarto estágio que coincide com o ano letivo de 2013/2014, no qual a diminuição do número de TI é acompanhada por uma diminuição no número de disciplinas; e (v) uma quinta fase, em que a oferta se resume a uma disciplina com um único teste.

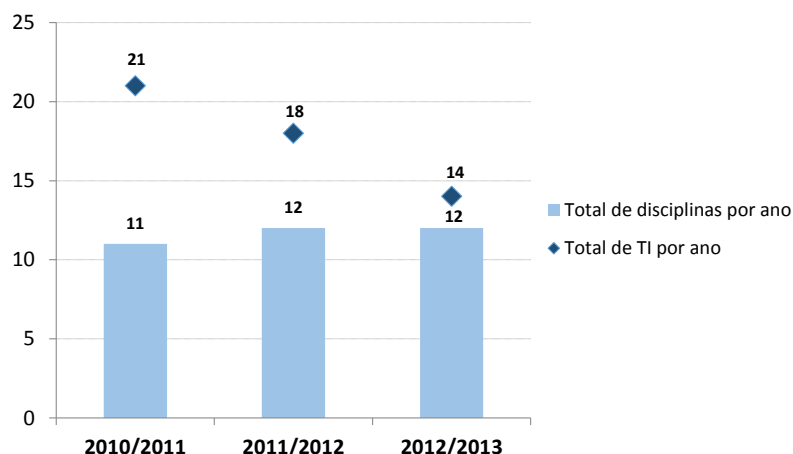
Gráfico 1 – Oferta de número de disciplinas e de TI, por ano (2005-2015)



O triénio que corresponde aos anos letivos 2010/2011, 2011/2012 e 2012/2013 constitui temporalmente a fase de evolução do Projeto TI que se estudou no presente trabalho. Para este efeito, foram solicitadas ao IAVE, I.P. as bases de dados de inscrições relativas àquele período, conforme adiante explicitado no capítulo relativo à metodologia.

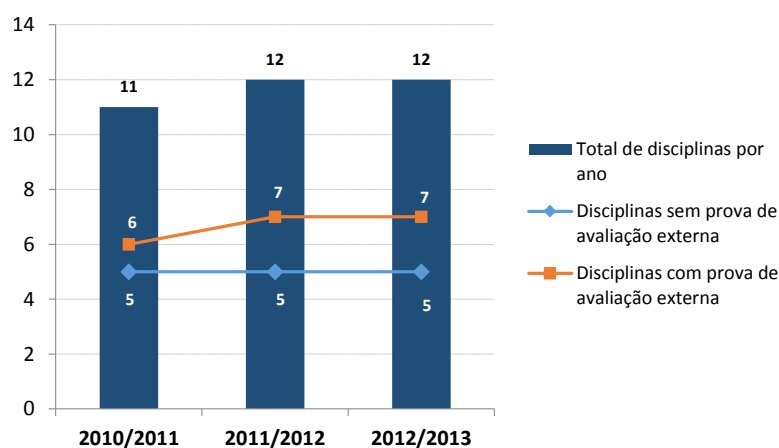
Na sequência do gráfico anterior, o Gráfico 2 destaca a oferta do número de disciplinas e de TI no triénio em análise, revelando uma descida acentuada do número de TI oferecidos (menos 1/3 entre o primeiro e o último ano).

Gráfico 2 – Oferta de número de disciplinas e de TI, por ano (2010-2013)



O Gráfico 3 identifica, no elenco de oferta anual, o número de disciplinas que, apresentando teste intermédio, têm provas de avaliação externa (provas finais ou exames nacionais). Nesta desagregação, fica evidente que, nos três anos, um pouco mais de metade das disciplinas oferecidas em TI são disciplinas que integram o elenco das provas de avaliação externa, mantendo-se constante, no triénio, a oferta de disciplinas sem provas de avaliação externa.

Gráfico 3 – Número de disciplinas no Projeto TI com avaliação externa e sem avaliação externa, por ano (2010-2013)



O Quadro 2 lista o elenco de testes, por ciclo de escolaridade, em cada ano do triénio 2010-2013. O conjunto é composto por 53 testes (21 do 3.º CEB e 14 do ES), distribuídos por 12

disciplinas diferentes (sete do 3.º CEB e cinco do ES). Os testes do 3.º CEB concentram-se no 8.º e no 9.º ano e os do ES distribuem-se pelos três anos do ciclo.

Uma caracterização pormenorizada do elenco do triénio, como a que é apresentada no Quadro 1, permite verificar que foram introduzidos novos testes de sete disciplinas (cinco do 3.º CEB e dois do ES), todos eles apenas com oferta de uma prova anual (ao contrário do que aconteceu em 2010/2011, por exemplo, com as disciplinas de Matemática – 9.º ano e de Matemática A – 12.º ano, que apresentavam duas aplicações anuais – t1 e t2).

A evolução ao longo dos três anos permite concluir que, enquanto na introdução das novas disciplinas só há oferta de um teste, as disciplinas que chegaram ao triénio já com edições anteriores vão perdendo testes, ficando progressivamente limitados a uma prova por ano (à exceção da disciplina de Matemática A – 12.º ano, que nos três anos mantém estável a oferta de duas provas por ano).

Quadro 2 – Elenco de TI disponibilizados (disciplina e ano de escolaridade), por ciclo de escolaridade (3.º CEB e ES), por ano (2010-2013)

Ciclos de escolaridade	2010/2011	2011/2012	2012/2013	No triénio
3.º CEB	CFQ_9*	CFQ_9	FQ_9	
	CN_9*	CN_9	CN_9	
	Geo_9*	Geo_9	Geo_9	
	Hist_9*	Hist_9	Hist_9	
	Ing_9*	Ing_9	Ing_9	
	LP_9	LP_9	Port_9	
	Mat_8	Mat_8		
	Mat_9_t1			
	Mat_9_t2	Mat_9	Mat_9	
N.º TI	9	8	7	24
N.º Disciplinas	7	7	7	7
ES	BG_10	BG_10		
	BG_11_t1	BG_11	BG_11	
	BG_11_t2			
	Filos_10*			
		Filos_11	Filos_11	
	FQ A_10	FQ A_10		
	FQ A_11_t1	FQ A_11	FQ A_11	
	FQ A_11_t2			
	Mat A_10	Mat A_10		
	Mat A_11_t1	Mat A_11	Mat A_11	
	Mat A_11_t2			
	Mat A_12_t1	Mat A_12	Mat A_12	
	Mat A_12_t2	Mat A_12	Mat A_12	
		Port_12*	Port_12	
N.º TI	12	10	7	29
N.º Disciplinas	4	5	5	5
N.º Total TI	21	18	14	53
N.º Total Disciplinas	11	12	12	12

Nota: Os testes das disciplinas assinaladas (*) são primeiras edições, isto é, foram realizados pela primeira vez nos anos letivos assinalados.

A escolha desta série temporal resultou de um conjunto de razões que se prendem com a constatação de que entre 2010/2011 e 2012/2013 se registou uma reconfiguração da oferta, resultante fundamentalmente de ter havido, no primeiro ano do triénio, uma evidente diversificação da oferta disciplinar (o número de disciplinas com TI praticamente duplica relativamente a 2009/2010), de o aumento de oferta disciplinar se ter feito quase exclusivamente à custa da introdução de disciplinas do ensino básico (há sete disciplinas do 3.º CEB em cada um dos anos do triénio, enquanto em 2009/2010 estas eram apenas duas – Matemática e Português) e, ainda, de o número de disciplinas e o elenco de disciplinas que

compõem a oferta se manterem estáveis, embora com uma diminuição progressiva do número de testes disponibilizados.

O ano de 2009/2010 é cumulativamente o ano de introdução de testes intermédios dirigidos ao 2.º ano de escolaridade, ciclo pela primeira vez abrangido pelo Projeto, conforme referido anteriormente. Reitera-se que o presente trabalho não se ocupará desta componente do Projeto, que constitui um traço distintivo do ano de 2010/2011, uma vez que a intervenção dirigida ao 2.º ano é enquadrada por finalidades específicas, que não são absolutamente sobrepostas às enunciadas para o 3.º CEB e ES, e tem, do ponto de vista administrativo, um conjunto de procedimentos distintos (de que decorrem, entre outros aspetos, períodos de inscrições e regras diferentes).

Às questões identificadas anteriormente como distintivas do triénio 2010-2013, todas de carácter substantivo, porque intrinsecamente relacionadas com a natureza central do Projeto enquanto intervenção de política educativa¹⁷, juntam-se outras de natureza mais logística e administrativa que suportam, complementarmente, a opção por este triénio, no histórico do Projeto TI.

De facto, o triénio em análise é o período em que se operam alterações nos procedimentos de adesão anual ao Projeto, nomeadamente no que se refere à forma de inscrição e também à metodologia de devolução de resultados. Em 2010/2011, o Projeto inaugura um processo de inscrição *online*, suportada por uma plataforma eletrónica, abandonando-se o processo de inscrição e de circulação de informação até aí usado (comunicações via correio eletrónico). Na referida plataforma, passam a ser executados todos os procedimentos que suportam, do ponto de vista logístico, as operações necessárias à implementação de um processo nacional desta dimensão: o acesso à documentação de enquadramento dirigida às escolas, a realização da inscrição, a escolha do elenco de testes a aplicar, o *download* dos enunciados dos testes (de acordo com o calendário de aplicação), a devolução de resultados, através do *upload* de grelhas de classificação, e o preenchimento de questionários de avaliação de cada um dos

¹⁷ Veja-se o enquadramento que a este propósito se apresenta no capítulo relativo à revisão da literatura, nomeadamente no que respeita à proposta de integração do Projeto TI no que se consideram ser modalidades de intervenção de políticas públicas relativas às áreas curriculares e pedagógica e à avaliação, informação e controlo.

testes e do Projeto na sua globalidade, entre outros procedimentos. É também neste ano letivo que, pela primeira vez, se assiste ao anúncio da denominada validação das inscrições, que se operacionaliza no ano seguinte (2011/2012). É um processo em que se faz depender o acesso à inscrição em cada TI da devolução, no ano imediatamente anterior, dos resultados do teste intermédio correspondente.

Não cabendo no âmbito do presente trabalho outras considerações sobre estas matérias específicas, não se pode deixar de as referir porquanto se considera que, em complemento das questões intrinsecamente relacionadas com o desenho da oferta avaliativa, estas questões de teor mais administrativo ajudam a delimitar um triénio marcado por uma clara evolução do Projeto TI, que nesta secção se define e caracteriza.

Em 17 de setembro de 2014, o IAVE, I.P. publicou, na sua página eletrónica a Informação n.º 1 – TI 2014/2015¹⁸, em que anunciou a não continuidade, nos *moldes vigentes*, das aplicações do 3.º CEB e do ES, no âmbito do Projeto TI.

Reproduzem-se, pela relevância para o presente trabalho, dois dos parágrafos da referida Informação:

Em relação aos testes aplicados no 3.º ciclo do ensino básico e no ensino secundário, considerando as evidências que os relatórios de anos anteriores têm revelado, nomeadamente o facto de ser residual, ou mesmo impercetível, o seu impacto na melhoria do desempenho dos alunos, entende-se, passados quase 10 anos desde o início do projeto, já não se justificar a sua continuidade nos moldes vigentes.

Atendendo ao vasto espólio de itens disponíveis no Banco de Exames e Provas, considera-se que fica assegurada a possibilidade de estes itens serem usados na familiarização com as provas de avaliação externa ou na autoavaliação dos alunos, sendo a publicitação anual de novos testes uma ação redundante e que pouco acrescenta à informação já existente (IAVE, 2014, p. 1).

¹⁸ Documento disponível para diretores, gestores de projeto e professores, através da Extranet do IAVE, I.P.. Documento reproduzido no Anexo D.

A opção pela descontinuidade da oferta de testes intermédios do 3.º CEB e do ES¹⁹ é comunicada às escolas já após o início do ano letivo 2014/2015 e parece querer configurar, nos termos em que é comunicada, uma decisão apoiada em evidências, em conhecimento empírico. Alega-se, com base nos relatórios dos anos anteriores, “o facto de ser residual, ou mesmo impercetível, o seu impacto [da aplicação dos testes] na melhoria do desempenho dos alunos”. Acrescenta-se que se “entende, por isso, passados quase 10 anos desde o início do projeto, já não se justificar a sua continuidade nos moldes vigentes” (IAVE, 2014, p. 1).

Já para o ano letivo 2013/2014, o elenco de TI proposto às escolas tinha sido consideravelmente reduzido, relativamente aos anos letivos anteriores. De facto, as alterações introduzidas nesse ano tornavam legítima uma perceção de reconfiguração do programa. Da leitura dos documentos relativos ao Projeto (na sua edição 2013/2014), não poderá resultar mais do que uma perceção, uma vez que explicitamente carecem de outro enquadramento ou definição capazes de sustentar as alterações realizadas.

Analisemos sinteticamente o teor da denominada Informação-Projeto – TI 2013/2014 (IAVE, 2013a), nomeadamente o seu ponto 1 – Apresentação e Finalidades do Projeto²⁰. Neste ponto do referido documento, são recuperados o ano inicial do Projeto (2005/2006) e o facto de primordialmente a intervenção se ter dirigido à disciplina de Matemática do 12.º ano, com o objetivo de melhor preparar os alunos para os exames, expondo-os a instrumentos de avaliação produzidos centralmente e de aplicação nacional síncrona (os TI). Rigorosamente, afirma-se a este propósito que a aplicação dos TI teve como primeira finalidade contribuir para a progressiva familiarização de todos os intervenientes com os instrumentos de avaliação sumativa externa, e que foi nesse sentido que se estendeu a oferta dos TI a outras disciplinas sujeitas a exame nacional. Nesta mesma informação (2013/2014),

¹⁹ A exceção a esta disposição é explicitada e justificada pelo IAVE no mesmo documento:

Constitui uma exceção à opção atrás referida a oferta do teste intermédio de Inglês (nível de continuação), a aplicar no 11.º ano de escolaridade, nos moldes adotados no ano letivo anterior. A aplicação deste teste, aferido para o nível B2 do Quadro Europeu Comum de Referência (QECR), enquadra-se numa estratégia de progressiva familiarização de alunos e professores com testes cujo nível de proficiência linguística de referência visa antecipar a prevista adoção, no nosso sistema de ensino, de um teste *First for Schools* (FCE). A aplicação deste teste será feita em contexto similar ao observado em 2014 com a aplicação do teste *Key for Schools* e do que se irá verificar, no corrente ano letivo, com a aplicação do teste *Preliminary for Schools* (PET), no 9.º ano de escolaridade. (IAVE, 2014, p.1).

²⁰ Documento disponível para diretores, gestores de projeto e professores, através da Extranet do IAVE, I.P.. Ponto 1 do documento reproduzido no Anexo E.

afirma-se que o alargamento da oferta dos TI a disciplinas não sujeitas a prova final de ciclo/exame final nacional teve na sua origem

o reconhecimento de que os TI podem ter um carácter eminentemente formativo — na medida em que, por um lado, permitem aos alunos a consciencialização da progressão da sua aprendizagem e, por outro lado, permitem aos professores a regulação das suas práticas, tendo por referência padrões de desempenho de âmbito nacional, mediante uma reflexão sustentada pela análise do processo de resposta dos alunos e pelos resultados atingidos (IAVE, 2013a, 1).

Ora, o carácter prioritariamente formativo dos TI, com funções de (auto)regulação das aprendizagens é assumido e anunciado como traço distintivo do Projeto em momento anterior ao do alargamento da oferta a disciplinas não sujeitas a prova final de ciclo/exame final nacional. De facto, já em 2008/2009, estas eram as finalidades de base, sendo que, na Informação n.º 1 desse ano letivo, se anunciava que a essas principais finalidades se juntava “por acréscimo” a de “progressiva familiarização com instrumentos de avaliação externa” (GAVE, 2008, p. 1). Só em 2010/2011 há de facto coincidência entre finalidades ou funções de carácter formativo e oferta de disciplinas não sujeitas a provas ou exames finais nacionais.

Serviu a análise sintética que se apresentou para sustentar aquilo que se afirmou não poderem ser mais do que perceções sobre a evolução que o Projeto sofreu em 2013/2014. De facto, o que serve de enquadramento à decisão de reconfiguração do Projeto neste ano (2013/2014) parece constituir um argumento de conveniência, mais do que uma explicitação lógica de evolução do Projeto.

De facto, só uma análise comparada do conteúdo dos documentos anuais de enquadramento do Projeto permite fazer uma aproximação àquilo que terá sustentado as decisões de evolução até ao anúncio da sua descontinuidade nos moldes até então vigentes. Em 2014/2015, começa por se abandonar a opção por um elenco alargado de testes dirigidos ao 3.º CEB e ES, assumidamente descontinuando a matriz do Projeto. Em fevereiro de 2015, cinco meses após o anúncio da não continuidade, não está publicado o relatório nacional do Projeto²¹ (publicação que desde 2008/2009 apresentava anualmente o balanço da

²¹ Há dois relatórios dirigidos exclusivamente aos TI do 1.º ciclo de escolaridade (2.º ano) – Relatório 2012/2013 e 1.º CEB - 2014, disponíveis em <http://iave.pt/np4/documentos?tema=36>

intervenção), pelo que há uma real carência de informação sobre outras razões relevantes para o rumo anunciado. Em rigor, o enquadramento dado à decisão carece de uma base empírica devidamente sustentada, que permitisse uma explicitação lógica da evolução do Projeto.

No último relatório nacional publicado até à presente data²², lê-se na introdução:

Na continuação do projeto, reitera-se o quadro conceptual e as finalidades já há muito explicitadas pelo GAVE: o valor de diagnóstico [com referência aos testes do 2.º ano explicitada em nota de rodapé] e o valor formativo e regulador destes instrumentos de avaliação, sustentados por uma ampla e detalhada informação sobre resultados. A informação por item e por aluno deve, pois, constituir, desde que bem utilizada e comunicada aos principais interessados, alunos e famílias, uma ferramenta quer de diagnóstico da qualidade da aprendizagem desenvolvida até então, quer de autoavaliação, para o aluno, quer ainda de possível instrumento de regulação, correção ou reajustamento das práticas dos professores. (GAVE,2013, p. 8).

Serve esta última referência para reiterar o quadro conceptual que até 2012/2013 (e pelo menos desde 2008/2009) servia de base ao Projeto²³, questão que nesta fase se convocou também porque é fundamental para a caracterização do objeto de estudo.

²² O referido relatório (relativo ao desenvolvimento do Projeto em 2011/2012) foi publicado em fevereiro de 2013.

²³ As referidas informações exigiriam, naturalmente, outras perspetivas de análise do seu conteúdo, designadamente, do teor da justificação apresentada no documento de setembro de 2014, a saber: *as evidências que os relatórios de anos anteriores têm revelado, nomeadamente o facto de ser residual, ou mesmo impercetível, o seu impacto na melhoria do desempenho dos alunos* (IAVE, 2014, p. 1). Não cabe, no âmbito do presente trabalho, a análise do Projeto centrada na medida da sua contribuição para a melhoria do desempenho dos alunos. Esta é uma matéria cuja avaliação é muito complexa e cuja abordagem terá de, certamente, ir além do que aquela que os mencionados *relatórios de anos anteriores* (apenas de 2009 a 2012) sistematizam.

CAPÍTULO 3 – REVISÃO DA LITERATURA

O enquadramento do objeto de estudo exige que se clarifique um conjunto de conceitos diretamente relacionados com dois domínios no âmbito da avaliação: *Avaliação em Educação e Avaliação no domínio das aprendizagens – avaliação interna e avaliação externa.*

AVALIAÇÃO EM EDUCAÇÃO

Há uma natureza eminentemente humana no ato de avaliar – discernir qualidade é um comportamento humano (Stake & Schwandt, 2006). Este seu carácter humano, a par da sua abrangência transversal e da omnipresença do processo avaliativo, revela a sua centralidade enquanto prática social. Dos programas escolares à área da saúde, da justiça aos demais outros domínios relacionados com a vida em sociedade, todos se podem constituir como objetos de avaliação, como domínios sobre os quais se podem produzir juízos de mérito e de valor. E são juízos formulados para servir a sociedade (ao serviço da sociedade), uma vez que providenciam representações de qualidade sobre o objeto avaliado, sustentando a melhoria (como e quando deve acontecer), a acreditação, a prestação de contas e quando necessário a fundamentação para pôr fim a programas cujos processos ou resultados foram considerados sem qualidade (Stufflebeam & Shinkfield, 2007).

O processo avaliativo envolve, naquela perspetiva, e de forma evidente, interação e responsabilidade sociais. É um processo que exige dos seus utilizadores a capacidade de se comprometerem com a utilização dos resultados, de lidarem com a representação de qualidade conseguida por um determinado processo. Avaliar é, assim, uma prática social com o objetivo de contribuir para resolver problemas ou melhorar aspetos das diversas áreas de uma sociedade (Fernandes, 2010).

Com este enquadramento, a definição de avaliação como *juízo sistemático do mérito e do valor de um objecto* (Joint Committee on Standards for Educational Evaluation – JCSEE, 2011) não satisfaz, por insuficiente. Como processo socialmente comprometido e de alcance tão diverso, necessita de incluir, além da representação do mérito e do valor, o juízo sobre a integridade (*probity*), exequibilidade (*feasibility*), segurança, significância (*significance*) e/ou equidade (*equity*) do objeto avaliado (Stufflebeam & Shinkfield, 2007). São critérios que contribuem para que o processo gere *afirmações avaliativas* mais sustentadas, credíveis e úteis (Fernandes, 2010).

E a forma como se faz e se utiliza a avaliação denuncia os pressupostos que enformam o paradigma de análise. Mais: revelam um estágio de evolução do paradigma. A existência de duas correntes, uma de pressupostos empírico-racionalistas (de pendor mais técnico) e outra assente em racionalidades interpretativas e sociocríticas, resulta numa dicotomia, que atualmente suporta a exigência de complementaridade. É a assunção e a defesa de que não há outro caminho do que aquele em que se sabe que “You have to use some of both.” (Stake, 2006, p xii). O que conjuga o pensamento criterial (a avaliação como medida) com o pensamento associado à avaliação como *experiência vivida* (Fernandes, 2010). Com o conhecimento proveniente da relação com as coisas e com os outros.

De facto, na base desta abordagem a *o que é avaliar* está a necessidade de assumir plena e cabalmente a afirmação de que discernir qualidade é um comportamento humano, de que a escolha de critérios, de *standards* de qualidade, é feita mediante intervenção de avaliadores, de que a avaliação necessita de interpretação/mediação dos avaliadores (no exercício de comparação, por referência a critérios, por referência a arquétipos, num exercício que é mais formal ou mais descritivo). Afinal, é partir do pressuposto de que o julgamento criterial tem raízes na interpretação de experiências pessoais (Stake & Schwandt, 2006). E que, por isso mesmo, a abordagem tem de ter cuidados acrescidos, tem de optar pelos mecanismos que mitiguem a dependência da acuidade e da credibilidade do avaliador, tem de saber partilhar a descrição, num processo de julgamento coletivo em que a experiência não é garante de interpretação correta, mas a inexperiência é quase a garantia de má interpretação (Stake & Schwandt, 2006, p. 411, tradução nossa).

A ponderação atribuída a cada uma daquelas vertentes, ou o peso que cada uma daquelas correntes assume (corrente empírico-racionalista e corrente interpretativa e sociocrítica), dependerá sempre das finalidades da avaliação e resultará sempre das características do objeto de avaliação e das condições de exequibilidade do processo avaliativo. Stake & Schwandt (2006) tratam estes critérios denominando-os como o carácter *embraceable* de uma avaliação, e para a sua definição concorrem a dimensão, a concentração e a distância do objeto a avaliar. É, na prática, deixar que a (im)possibilidade de haver contacto direto e continuado com o que se está a avaliar decida do peso relativo que a avaliação de base criterial ou de base experiencial assumem no processo. E daqui não decorre uma escolha por uma ou por outra abordagem, decorre, sim, a consciência de uma subsidiariedade, em que as abordagens se complementam, posicionando-se no plano do que é exequível. Não se excluem metodologias quantitativas (que Stake orgulhosamente lista no seu *Standards-based & responsive evaluation* – instrumentos quantitativos, *standards*, parâmetros de pontuação/classificação e outros instrumentos criteriais), faz-se, antes, a defesa de que é fundamental apostar numa abordagem de compromisso²⁴.

E o que se defende é que outra coisa não será possível, porque, independentemente da metodologia aplicada, a representação da qualidade é sempre um ato de seleção, é sempre uma sobressimplificação, e o que é reportado é um exercício de interpretação. Portanto, resultado de discernimento humano.

E reconhecer aquele *óbvio*, saber trabalhar com ele, é uma lógica de abordagem que se recupera dos ensinamentos que o reputado JCSEE tem vindo a apresentar nas atualizações das suas publicações sobre *standards* de avaliação de programas. A propósito de outros enunciados de carácter normativo ou referencial relativos à conceção e implementação de avaliações de programas, o Comité investe frequentemente em (re)afirmar que é fundamental saber reconhecer, saber trabalhar com e saber gerir aquilo que muitas vezes

²⁴ Na explicitação da denominada avaliação responsiva (Stake, 2006), recupera-se a recusa do tratamento dicotómico. De facto, tal como na abordagem da avaliação baseada em *standards*, também nesta perspetiva convivem necessariamente as lógicas criteriais e interpretativas. A diferença é que o protagonismo é dado à observação interpretativa, à predisposição experiencial no conhecimento global do objeto em avaliação. E esta predisposição, defende Stake, pode ser levada a cabo em qualquer avaliação. Afinal, é a que melhor configurará a vontade de descobrir “what’s out there” (p. 86).

apresenta natureza facilmente apreensível, mas que, de tão supostamente admissível e consensual, ou, então, de só aparentemente universal, é frequentemente desvalorizado ou esquecido.

Na conceção de processos de avaliação de programas/projetos, a finalidade será a de definir a perspetiva em contexto. Fernandes (2010) identifica três elementos principais que deverão estar presentes: a teoria do programa; o rigor de métodos e procedimentos; a relação estrita e direta com a natureza das questões e a utilidade do tipo de respostas a providenciar. É uma abordagem em que o conhecimento do objeto de avaliação, do contexto e dos propósitos da avaliação pauta as escolhas: quais são as questões de avaliação no quadro da intervenção que o programa/projeto constitui ou pretende constituir; que métodos e técnicas de recolha de dados respondem de forma mais fiável e rigorosa às questões colocadas; que respostas se conseguirão alcançar; que recomendações se conseguirão traçar com os métodos convocados, de forma a gerar informação útil para o destino a dar ao programa/projeto? Em suma, são os propósitos da avaliação que ditam a abordagem e a utilidade é um elemento importante na condução de todo o processo.

De acordo com Stufflebeam & Shinkfield (2007), é o propósito da avaliação que preside, entre outros, à escolha da modalidade: comparativa/não comparativa – o desenvolvimento qualitativo de um produto dispensa uma avaliação por comparação, mas a confirmação da evolução para a maturidade desse produto exigirá uma abordagem comparativa.

Aquela dependência entre a finalidade da intervenção, os seus usos, e o carácter que a avaliação assume é evidente na distinção entre avaliação formativa e sumativa. A primeira intimamente ligada à busca da melhoria, de natureza processual, prospetiva e proativa, e a segunda com o intuito de produzir informação para o apuramento e a atribuição de responsabilidades e para a prestação de contas, baseada num juízo de orientação retroativa (podendo ter por base informativa a avaliação formativa). Aos usos anteriores, os autores juntam o de disseminação (muito dirigido a suportar escolhas de consumidores) e o de esclarecimento, ou de produção de conhecimento (que, por regra, conduz a processos de investigação).

No terreno, encontramos *designs* avaliativos que congregam finalidades de prestação de contas, de regulação e de potenciação de resultados e mesmo de produção de conhecimento. E, mais uma vez, são os propósitos da avaliação e o contexto a pautarem a ação do avaliador. É o enquadramento conceptual a informar escolhas que se querem conscientes e informadas, mas a gerarem respostas sustentadas e úteis. É idêntico ao que Fitzpatrick (2009) conclui na primeira análise global que apresenta na obra que reúne doze entrevistas realizadas a conceituados avaliadores sobre processos de avaliação implementados sob a coordenação de cada um deles. O objetivo central desta análise ao *corpus* de informação recolhido pelas entrevistas é “restaurar a centralidade da prática na avaliação” (Fitzpatrick, 2009, p. 355, citando Schwandt, 2005). Na sua análise, a autora pretende que as escolhas feitas por cada um dos avaliadores falem por si, enquanto evidências de opções conscientes, tomadas em contexto real, no contexto de como se faz avaliação. É uma lógica reconhecível no *discernimento pragmático* que Fernandes (2010) reclama para a prática da abordagem avaliativa: a prática é apresentada como meio de conhecimento, uma vez que tem natureza processual e deliberativa (revelará o que é tácito se devidamente interpretada).

A intervenção política (entendida como ciência do governo das nações) na área da educação constitui naturalmente um domínio em que a avaliação é uma exigência. A exigência resulta, simplificadamente, da necessidade absoluta de prestar contas e da não menos importante necessidade de produzir informação e conhecimento capazes de melhorar a intervenção política (Rodrigues, 2010), na busca de fundamentação para as decisões de governo.

Rodrigues (2010) distingue seis modalidades de intervenção no âmbito das políticas educativas: (i) medidas de distribuição ou afetação de recursos humanos, financeiros ou tecnológicos essenciais à prestação do serviço de educação; (ii) organização, gestão e administração dos recursos públicos; (iii) organização curricular, de programas de ensino e de orientação pedagógica; (iv) matérias relativas à produção de informação e conhecimento, à avaliação, acompanhamento e controlo dos vários elementos constitutivos do sistema educativo; (v) medidas de política simbólicas que visam distinguir, sublinhar ou reconhecer a qualidade ou excelência; e (vi) medidas de política constituintes.

Atendendo ao objeto do presente trabalho, importa considerar as modalidades relativas a (i) áreas curricular e pedagógica e (ii) avaliação, informação e controlo. As primeiras incluem as decisões relativas ao conteúdo e às metodologias de ensino, sendo a segunda modalidade a que enquadra as ações no âmbito do que é reconhecido como instrumentos de avaliação externa das aprendizagens (exames e provas de aferição), enquanto mecanismos de produção de informação e conhecimento sobre o sistema²⁵.

Para a avaliação de uma determinada intervenção de política pública, nomeadamente de um programa de política pública, será preciso conhecer a teoria do programa, isto é, o enquadramento que nos ajuda a determinar de que forma um projeto, programa ou intervenção visa resolver um dado problema social (Fernandes, 2010). Numa avaliação em que se tem em conta a *teoria do programa*, tem de ser considerado o processo através do qual se espera que os componentes do programa afetem os resultados, bem como as condições em que se acredita que os processos operam (Donaldson, 2001). O conhecimento da teoria de um programa revela o seu funcionamento, organização, relações entre componentes, processos utilizados e efeitos pretendidos. A sua análise informa o desenho da avaliação, uma vez que permite identificar e selecionar questões relevantes, definir procedimentos adequados de recolha de informação e identificar dimensões do programa que devem merecer particular atenção, porque serão as mais críticas para o seu sucesso (Fernandes, 2010).

É exatamente partindo do pressuposto de que a intervenção de avaliação de um programa, projeto ou sistema é pautada pela melhoria que a OCDE desenvolve o projeto *OECD Review*

²⁵ A relação destas duas áreas de modalidades com o Projeto TI reside no facto de, a partir do seu enquadramento conceptual, se reconhecer aos instrumentos de avaliação disponibilizados (os TI) um carácter de avaliação, de base, externa, porque concebida centralmente, e de aplicação síncrona e processualmente controlada, embora se reconheça, igualmente, que os TI apresentam uma natureza híbrida, uma vez que a realização e classificação seguem práticas próximas das usadas na avaliação de carácter interno, e não das que distinguem a avaliação externa (no caso dos TI, a aplicação pode ser feita na turma, com professor da turma, em tempo de aulas, com classificação assegurada pelos professores dos alunos). Complementarmente, as finalidades que enquadram o Projeto, a par das características que o aproximam de intervenções de avaliação externa, configuram os TI como instrumentos que, se devidamente apropriados, contribuem para a regulação das aprendizagens, para o desenvolvimento de trabalho colaborativo, para a deteção atempada de dificuldades, para a geração de informação avaliativa de qualidade (no contexto de cada escola e, também, no quadro de padrões nacionais). Todas estas características perfilam-nos como capazes de contribuir para o apuramento de conteúdos curriculares e de metodologias de ensino.

*on Evaluation and Assessment Frameworks for Improving School Outcomes*²⁶. O propósito é o de melhorar a qualidade, equidade e eficiência da educação escolar, olhando/analizando as várias componentes dos quadros de referência de avaliação (de aprendizagens, de professores, de escolas e de sistemas) usados pelos diversos países com o objetivo de melhorar o desempenho dos alunos.

Neste âmbito, a OCDE, através da sua Direção de Educação, identifica cinco elementos-chave para análise:

- (i) Designing a systemic framework for evaluation and assessment; (ii) ensuring the effectiveness of evaluation and assessment procedures; (iii) developing competencies for evaluation and for using feedback; (iv) making the best use of evaluation results; and (v) implementing evaluation and assessment policies (Santiago *et al.*, 2013, p. 657).

Nesta fase, interessará fundamentalmente o domínio do *design* de um quadro sistémico de avaliação, porquanto é na assunção de que o conhecimento do sistema no seu conjunto, da sua organização, da relação entre as partes, e dos objetivos pretendidos que residem as orientações para a planificação e implementação do processo de avaliação. Não haverá forma de poder (bem) avaliar um sistema, ou programa, ou projeto, se não se conhecer o seu enquadramento de base, a sua teoria.

E, no quadro das evoluções que Portugal protagonizou nos últimos anos no âmbito da política educativa (designadamente na área da avaliação), conforme já se referiu em momento anterior, uma das recomendações constantes no Relatório sobre Portugal (Santiago *et al.*, 2012) reitera aquela necessidade de se consolidar o *framework* nacional de avaliação (estabelecendo um racional claro e uma narrativa convincente sobre como a avaliação se relaciona com o programa de reforma do sistema educativo, descrevendo como cada componente do *framework* avaliativo poderá produzir resultados úteis para as práticas de sala de aula e para a generalidade das práticas escolares de melhoria).

Serve esta referência para enfatizar a necessidade de consolidação dos referenciais das políticas de intervenção pública, condição sem a qual as avaliações tenderão a ser menos

²⁶ Informação disponível em www.oecd.org/edu/evaluationpolicy

bem conseguidas. Afinal, se avaliamos o que conhecemos menos bem, também porque os limites do objeto de estudo não são claros, ou não estão consolidados, não teremos, à partida, as melhores condições para, com base na teoria que sustentará as intervenções, identificar e seleccionar as questões relevantes, as que, no âmbito da avaliação em educação, contribuiriam para formular juízos que permitiriam a melhoria.

AVALIAÇÃO NO DOMÍNIO DAS APRENDIZAGENS: AVALIAÇÃO INTERNA

No domínio da avaliação das aprendizagens, partir da perspectiva *Avaliar para melhorar* não é excluir as representações que lhe são prévias (*Avaliar para medir; Avaliar para descrever, Avaliar para ajuizar*). Guba e Lincoln (1989) assumem claramente que cada uma destas etapas representou um *passo em frente*, não só na qualidade e variedade de matéria convocada para o processo avaliativo, mas também no nível de sofisticação alcançado.

O caminho passará por reconhecer limitações aos modelos anteriores e procurar aquele que possa dar resposta aos desafios da diversidade que marcam a sociedade e a escola nas últimas décadas, e atender às alterações dos paradigmas de investigação, nos domínios das teorias da aprendizagem, do ensino e da avaliação.

Veja-se que existe uma estreita relação entre os paradigmas de investigação/avaliação e a abordagem que é defendida sobre como se pode aferir o que se aprendeu: uma conceção mais positivista encara as aprendizagens dos alunos como uma realidade objetiva, passível de ser medida e apreendida na sua totalidade através da aplicação de testes, que visam medir se os conhecimentos foram adquiridos ou se os objetivos previamente definidos foram alcançados (esta última finalidade a enquadrar uma perspectiva behaviourista). Esta é uma visão da aprendizagem (e portanto do ensino) como se se tratasse de um *building-block model* (Gipps, 2012), aprendizagem por pequenas unidades, linear, sequencial, numa lógica subordinada à indispensabilidade dos pré-requisitos.

Por outro lado, uma concepção cognitivista e construtivista encara a avaliação das aprendizagens como um processo complexo, subjetivo e que não nos permite apreender na totalidade o que os alunos sabem e são capazes de fazer. Nesta visão, a avaliação está intrinsecamente associada ao ensino e à aprendizagem, dá relevância aos contextos, à participação de todos os envolvidos (antes de mais, professor e alunos) e aos processos cognitivos e sociais envolvidos. Neste paradigma, o modelo de avaliação das aprendizagens já não pode *atomizar conhecimentos* (Gipps, 2012, p.18).

No âmbito específico da avaliação educacional, é incontornável recuperar Michel Scriven e a distinção que faz, em 1967, no contexto da avaliação de programas, entre avaliação formativa e sumativa. Fernandes (2005) convoca os conceitos, afirmando que, em contextos educativos, a modalidade formativa se destinava a ir corrigindo e ajustando o ensino e a aprendizagem e uma avaliação sumativa, no essencial, fazia um balanço e emitia um juízo final acerca do que os alunos sabiam e eram capazes de fazer.

A partir daquele quadro inicial, de que Scriven foi fundador, começou por prevalecer uma visão eminentemente behaviourista e neobehaviourista: a avaliação formativa é a que recolhe informação consubstanciada em produtos de aprendizagem, a medir por referência a objetivos comportamentalmente definidos, que, se não atingidos num tempo determinado, darão origem a tarefas de remediação. Nesta perspetiva, a informação tem valor retroativo, mede resultados, embora não o faça com frequência, e tem uma vocação preparatória da avaliação sumativa.

Ora, esta avaliação formativa é extrínseca ao processo de ensino e de aprendizagem. Acompanha-os, mas não os integra. O enfoque continua a ser o da contribuição para a atribuição de classificações, como um meio para chegar ao que se pretende: mais medir, descrever ou ajuizar, do que melhorar. É o próprio Scriven que em 1980 reelabora os conceitos, enfatizando que a avaliação formativa é a que é conduzida *portas para dentro*, para *o in-house staff of the program* e normalmente *fica em casa*. Usa, nesta sua reformulação, uma imagem de Stake, que reproduzimos no original: “When the cook tastes the soup, that’s formative; when the guests taste the soup, that’s summative.” (citado por Scriven, 1991, p. 19)

Conceção diferente é, de facto, esta em que a avaliação faz parte integrante do processo, da construção (da confeção, por proximidade à imagem de Stake) do que se ensina e se aprende. E a avaliação interna, centrada na escola, nos professores e nos alunos, encontra a sua primeira vocação na avaliação formativa. É a *avaliação para as aprendizagens*²⁷ – integradora, na medida em que relaciona estreitamente ensino, aprendizagem e avaliação, centrando-se no papel do professor e do(s) alunos(s) como atores de interações diversificadas na sala de aula. Será sempre este espaço de interação que merecerá enfoque, é a *black box* que exige conhecimento e avaliação continuada do seu complexo sistema interno: aquele em que se procede à organização do processo de ensino (das suas práticas), à definição dos seus propósitos, à seleção das estratégias, à regulação de estratégias através de *feedback*, à reorientação das práticas, e no qual se vão criando ambientes de aprendizagem através de comunicação efetiva. Concomitantemente, os alunos são chamados à participação ativa no seu processo de aprendizagem (e, naturalmente, de avaliação): potenciando o *feedback* recebido, analisando o seu trabalho, organizando e regulando o seu percurso e partilhando-o com os seus colegas.

A avaliação formativa, cujo racional tem vindo, como já se referiu, a evoluir nas últimas décadas, não deverá ter base “casuística ou intuitiva” (Santos, 2010). É um processo de acompanhamento e de regulação do ensino e da aprendizagem, cujo enfoque é o trabalho do dia a dia. É um processo em que o *feedback* assume papel central, porque o *feedback* representa a efetivação de que o processo educativo (que é, afinal, um processo comunicativo) pretende atingir verdadeira interação discursiva (entendendo-se que esta abrange a pragmática, o conhecimento do mundo, do meio social e cultural), i.e., que a área de intersecção entre os interlocutores (professor e aluno) está, por meio de *feedback* variado e de qualidade, a ser progressivamente alargada, porque se vai conseguindo saber mais, compreender melhor, o que contribui para a fluidez e para o sucesso da comunicação, ou seja, para o sucesso da aprendizagem.

²⁷ Assume-se a definição do Assessment Reform Group (ARG): “Assessment for learning (AfL) is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there” (2002, p. 2).

É no quadro de uma avaliação entendida como um processo que deve servir a elevação da aprendizagem que Butt (2010) também perspetiva a modalidade de avaliação sumativa, de natureza certificativa. Esta é a *avaliação das aprendizagens*²⁸ que produz juízos de valor, que faz interpretação de desempenhos e que agrega informação de natureza quantitativa, que permite fazer pontos de situação, fazer registo de desempenhos alcançados, classificar, divulgar informação aos alunos, e também a outros interessados, exteriores à sala de aula. Tem carácter não continuado, é pontual, pode desencadear processos retroativos ao ser usada como *feedback* para as planificações de ensino. A informação que gera pode ser usada para reconstruir percursos didáticos que tenham revelado insuficiências, bem como para devolver ao aluno mais do que uma classificação (cujo potencial informativo é naturalmente pobre).

Sendo a avaliação parte de todo o processo, as suas funções – formativa e sumativa, nas suas utilizações mais ou menos formais – devem ser claramente definidas nas suas características, potencialidades e constrangimentos, para que, no âmbito da avaliação interna, elas possam contribuir ambas para apoiar as aprendizagens, melhorando-as, com vista à elevação dos *standards*.

Sobre as relações entre a avaliação formativa e a sumativa, será importante destacar a perspetiva de Harlen (2006a), que aborda as duas naturezas da avaliação como pertencentes a um contínuo da dimensão avaliativa. Nesta perspetiva, os dados resultantes da avaliação formativa não deverão ser recolhidos e utilizados com objetivo de classificação, pois, dessa forma, o propósito primeiro desta avaliação (de que faz parte ser formativa, integrar o processo e ser reguladora) seria posto em causa. Daqui não decorre que a avaliação formativa tenha de ocupar posição dicotómica relativamente à sumativa, porque ambas pertencem a essa dimensão avaliativa, sem que haja descontinuidade entre as práticas de uma e de outra. De facto, nesta linha não descontinuada, há áreas de compatibilidade, sendo que, por exemplo, os dados recolhidos em avaliação sumativa podem ser usados para fins formativos (sendo o contrário também possível, embora tendencialmente mais problemático). Na prática, feito o julgamento, que proporciona uma súmula do que o aluno aprendeu, as

²⁸ Segue-se a definição do Assessment Reform Group (ARG): Assessment of learning “is used to summarise what pupils know or can do at certain times in order to report achievement and progress” (2003. p.2).

evidências do desempenho que se classificou revelarão, por exemplo, onde se concentram as dificuldades, permitindo ao professor projetar a intervenção seguinte de forma sustentada. Neste sentido, a avaliação sumativa não se esgota na mera atribuição de "notas".

Importa aqui sublinhar que não decorre da opção qualitativa vs quantitativa a essência da avaliação realizada (de facto, a forma como se recolhe informação não determina, à partida, o uso que dela se vai fazer para a aprendizagem – um teste realizado em aula pode servir para alimentar opções de ensino sem dele terem resultado classificações, ou pode fornecer uma informação do nível atingido relativamente a critérios definidos). Importa, igualmente, registar que de um uso duplo – (i) re-interpretação sumativa da avaliação formativa, i.e., uso sumativo dos dados recolhidos em avaliação formativa; e (ii) uso formativo da avaliação sumativa – poderá resultar uma preponderância da avaliação sumativa, por ser aparentemente mais eficiente. Equivale isto a dizer que, se, por hipótese, tivéssemos de escolher entre avaliação formativa e avaliação sumativa, esta última poderia parecer a mais capaz de responder a um uso duplo. Ora, sabendo que é a avaliação formativa a que privilegia a recolha detalhada e repetida de elementos, a que melhor conseguirá denotar a progressão na aprendizagem, e que é aquela que, embora marcada pelo enviesamento do olhar particular, casuístico, de um avaliador, é suscetível de ser corrigida pela recolha de outros elementos (e, portanto, detentora de maior qualidade), a primazia da avaliação sumativa implicaria a perda de elementos distintivos do processo que se pretende *para as aprendizagens*. Perder-se-ia: a noção de progressão alimentada por dados detalhados, recolhidos em momentos diversos, com interpretação suscetível de correção. Perder-se-ia, assim, a integridade da avaliação formativa, a única capaz de, no trabalho desenvolvido em aula, com o julgamento centrado no aluno, ser protagonizada por este e pelo professor, dando-lhes simultaneamente *feedback* para que cada um possa saber o que e como ensinar ou aprender.

Que a avaliação se constitui como um dos principais mecanismos para a promoção de aprendizagem efetiva é consensual. A declaração é feita pelo Assessment Reform Group (ARG) num dos seus vários trabalhos²⁹, na sequência do conhecido *Inside the black box*:

²⁹ Referimo-nos a *Assessment for Learning. Beyond the black box*. (ARG,1999, p. 2)

Raising standards through classroom assessment (Black & Wiliam, 1998). Mas os autores apressam-se a acrescentar que o *mecanismo* tem de ser usado de forma correta e que não há evidências de que o aumento de momentos de testagem tenham, por si, implicações positivas na aprendizagem. O desafio é mesmo o de conseguir que a avaliação estabeleça uma ligação muito estreita com o ensino e com a aprendizagem, que se foque no que acontece na sala de aula, na natureza das interações entre professores e alunos e na qualidade da sua experiência educativa (Butt, 2010). E o professor, enquanto *providenciador* de *feedback* (dando aos alunos informação de qualidade sobre os seus desempenhos) e enquanto *utilizador* de *feedback* (usando a informação avaliativa para corrigir, recentrar, sustentar as suas opções didáticas) é o protagonista desta interação. E enquanto elemento orientador, na sua interação com os alunos, é necessário que o professor tenha uma perceção muito clara dos objetivos, conhecimentos e processos que a tarefa de aprendizagem convoca, assim como dos propósitos da tarefa avaliativa, da forma ou instrumento a usar, do domínio do saber que é objeto de avaliação (entre outros aspetos que integram a planificação) – sem dominar a natureza do que exige, dificilmente poderá alimentar um processo de melhoria (Pinto & Santos, 2006).

Ora, se o professor desempenha o papel principal no processo, a sua perspetiva sobre o que a avaliação pode fazer pela melhoria da aprendizagem é determinante na sua ação educativa. Para que a ação seja consequente, é fundamental que a conceção de base do professor seja basicamente coincidente com a premissa de que *todos os alunos podem aprender*. Se assim não for, o professor não reconhecerá a avaliação como capaz de elevar *standards* de aprendizagem, uma vez que das suas convicções fará parte a ideia de que o que o aluno consegue aprender e como o faz estará em grande parte pré-determinado (dependente das suas capacidades intelectuais à partida) e, portanto, não é suscetível de ser sujeito a intervenção relevante.

Nessa visão, a avaliação, afirma Butt (2010), “é simplesmente um meio de medir (ou de reafirmar) o que as crianças estão geneticamente pré-determinadas a aprender, sendo que a avaliação terá pouco interesse educativo” (p. 50, tradução nossa). Essa será, afinal, a posição

de um professor não comprometido, porque acredita que o seu envolvimento e as suas práticas avaliativas não teriam repercussões relevantes no sucesso dos seus alunos.³⁰

Uma avaliação capaz de promover melhor aprendizagem é a que se constrói com professores “engaged”, é uma conceção em que se advoga como prioritário ajudar os professores a encararem a avaliação como parte integrante do ensino e da aprendizagem, “dando proeminência [a esta abordagem] na formação inicial de professores e na formação de desenvolvimento profissional” (ARG, 1999, p. 12, tradução nossa).

Desenvolver e implementar boas práticas de avaliação (não se distinguindo aqui propositadamente a natureza formativa da natureza sumativa), construir tarefas avaliativas diversificadas, indutoras de aprendizagem, desafiadoras e complexas, capazes de mobilizar e integrar saberes, e de serem consonantes com as tarefas de aprendizagem, e reflexo do que no currículo se considera estruturante, é, desta forma, fundamental.

Essa será a visão que abre caminho para que o *diálogo*³¹ entre a avaliação *para e das* aprendizagens possa ser feito com qualidade e bons resultados.

Será sempre um diálogo em que deverão ficar salvaguardadas as diferenças, definidas as áreas de intervenção, na medida em que a avaliação formativa e a sumativa servem propósitos diversos, embora concorrentes no que respeita à aprendizagem.

Afinal, o processo avaliativo necessita de ensino explícito, em que o professor sabe o que tem de ser aprendido, domina a tarefa pedida, o conjunto de critérios a seguir, o *standard* a atingir. Em síntese, o professor sabe ser explícito sobre os propósitos da aprendizagem, e,

³⁰ Esta perspetiva, aqui necessariamente abordada de forma sintética, mereceria enquadramento com outras questões, como, por exemplo, as relativas à motivação, ao contexto socioeconómico e cultural, à noção de escola enquanto valor acrescentado, entre outras aspetos. O que se pretende ressaltar é que outras questões concorrem para o sucesso do processo ensino-aprendizagem-avaliação, i.e., o professor que não adota uma perspetiva em que os alunos têm irremediavelmente capacidades pré-determinadas tem de contar com todos os outros desafios que resultam das características específicas dos alunos, da sua história pessoal, da sua proveniência, da forma como a escola enquanto instituição gere as diferenças, uma vez que algumas das competências centrais para o sucesso académico serão influenciadas por estes aspetos.

³¹ Referência ao título do Capítulo V – “*We need to talk*” (Butt, 2010).

por maioria de razão, sobre os propósitos da avaliação (sendo que para esta necessidade ser satisfeita concorre naturalmente o seu saber científico e didático³²).

Na prática, ambas se alimentam daquilo que um *feedback* de qualidade conseguir gerar: os alunos recebem resposta formal ao seu desempenho, comunicando-se em que medida é que o trabalho feito, a resposta dada, se relaciona com os critérios e como é que o desempenho poderá ser melhorado para atingir um determinado *standard* (na lógica de que o *feedback* é, como afirma Stobart (2008, p. 148), “to close the gap between where learners are at present and where they need to get to” – “the desired performance”). Na situação concreta da avaliação de natureza sumativa, é feito um juízo, num dado momento, sobre se o trabalho atingiu o *standard* pretendido. E para que estes dois propósitos sirvam a melhoria da aprendizagem têm necessariamente de ser coerentes e integradores. Sem se desvirtuarem, têm de ser prudentemente complementares. De outra forma, estaríamos a trabalhar para “ajudar a aprender” uma *coisa* diferente da que íamos avaliar. Seria profundamente incongruente. Não seria nem justo, nem transparente. Não contribuiria certamente para uma avaliação que se quer autêntica e fiável, em que os alunos devem ser capazes de perceber sem problemas a sua situação face às aprendizagens que têm de adquirir ou desenvolver. No fundo, é tão poderosamente simples quanto fazer coincidir as tarefas da avaliação com as da aprendizagem³³ (Fernandes, 2005).

³² Gipps (1994) afirma sobre esta matéria: “Teachers cannot assess well subject matter they don’t understand just as they cannot teach it well. Teachers have to understand the construct they are assessing (and therefore what sort of tasks to set).” (p. 135).

³³ No original: “fazendo coincidir as tarefas de aprendizagem com as de avaliação” (Fernandes, 2005, p. 80).

AVALIAÇÃO NO DOMÍNIO DAS APRENDIZAGENS: AVALIAÇÃO EXTERNA

A implementação de programas de avaliação externa, em larga escala, de âmbito nacional ou regional, através de provas estandardizadas, de conceção e classificação centralizadas, com aplicação em condições controladas, desde há muito que se constitui como um campo de ação das políticas públicas de educação, respondendo a uma diversidade de propósitos.

Fernandes (2014) denomina os objetivos de moderação das avaliações internas, de supervisão do desenvolvimento do currículo, de monitorização dos resultados de escolas e de alunos e de contribuição para a justiça e equidade da avaliação como *politicamente poderosos* e com larga base de sustentação social. Este apoio, que os decisores, de diferentes famílias políticas, recebem, resultará, em larga medida, da perceção de que os instrumentos usados em avaliação externa são medidas objetivas do desempenho dos alunos, das escolas, do trabalho dos professores, do sistema, proporcionando informação insuspeita. É principalmente um exercício com objetivos de prestação de contas, o que, entre vários outros aspetos, também pretende assegurar, do ponto de vista político, o valor das despesas realizadas pelos contribuintes (Madaus, Russell & Higgins, 2009).

Uma abordagem da temática da avaliação externa beneficiará da assunção de que é matéria dominada por uma controvérsia alimentada por ideias pré-concebidas e extremadas (Fernandes, 2014). As provas de avaliação externa são muitas vezes vistas como formas oficiais de discriminação social e mesmo de segregação, responsáveis por um conjunto extenso de problemas que afetam a qualidade do que se ensina e se aprende e o desenvolvimento do currículo (e.g., Esteban, 2009; Casassus, 2009; Pacheco, 2013). A dicotomia é construída com a avaliação interna a ser prejudicada, mesmo desvirtuada, pela existência de provas de avaliação externa.

Outra dicotomia se constrói nesta complexa tomada de posições: a que opõe os adversários dos exames aos que os identificam como uma *key strategy for improving educational quality*, *o alimento da reforma*, ou mesmo o *Maná* (Madaus, Russell & Higgins, 2009), um sistema de monitorização do desempenho dos alunos e um veículo de mudança, capaz de conduzir o

que se ensina, se aprende e como se ensina e se aprende, capaz de motivar os desmotivados, elevar *standards*, e guindar países para níveis superiores de produtividade e de competitividade.

No quadro do que foi afirmado, será fundamental analisar a questão de uma perspectiva conceptualmente informada.

Importará começar por partir do pressuposto (mesmo que por bondade do argumento) de que os programas de avaliação externa terão, independentemente dos propósitos que em cada momento ou contexto servem, uma finalidade global de melhoria da qualidade das aprendizagens, que significará melhor *performance* dos sistemas. A prossecução dessa finalidade requer, segundo Mehrens (2002), o desenvolvimento de um processo socialmente pilotado em que todos os envolvidos partilhem uma linguagem comum. Esta perspectiva faz da implementação de um sistema de avaliação externa uma abordagem socialmente alargada, envolvendo alunos, professores, pais, direções de escolas, responsáveis locais e nacionais. É a perspectiva que pretende tirar partido de um processo que, na base, terá a necessidade, incontornável, de os sistemas educativos serem prestadores de contas, capazes gerar resultados comparáveis num universo alargado.

Valerá a pena relembrar (ainda que aparentemente *fora do tempo*) as conclusões sobre como elevar os *standards* de aprendizagem a que Black e William (1998) chegaram no seu seminal trabalho. Efetivamente, o seu estudo revelou que (i) a prática sistemática de avaliação formativa melhora substancialmente a aprendizagem de todos os alunos, (ii) os alunos que revelam mais dificuldades de aprendizagem são os que mais beneficiam de tais práticas, (iii) os alunos que frequentam aulas em que a avaliação é predominantemente de natureza formativa obtêm melhores resultados em provas de avaliação externa.

Serve a recuperação das virtualidades da avaliação interna para que estas possam ser relacionadas com o que pode ser a vontade de tirar melhor partido da avaliação externa. Hamilton (2003) começa por reconhecer que uma e outra (avaliação interna e externa) exercem efeitos significativos nos alunos, para, de seguida, propor linhas de ação dirigidas à promoção de políticas de testagem em larga escala que sejam eficazes, isto é, que

contribuam positivamente para a elevação da qualidade do desempenho dos alunos. As linhas de ação propostas centram-se em três áreas, a saber: (i) investimento na melhoria do *design* das provas (com enfoque, entre outros temas, nas questões relacionadas com a validade); (ii) aposta na melhoria do *design* dos sistemas de prestação de contas (por exemplo, através da devolução de resultados aos professores de forma atempada e útil, do desenvolvimento de investigação cujo objeto é o alinhamento entre *standards*, avaliação, currículo e ensino, e, ainda, através da formação de professores; e (iii) implementação de uma intersecção entre o sistema de avaliação interna e o de avaliação externa, isto é, promovendo o alinhamento entre o trabalho de aula e os *standards* estabelecidos pelas provas de avaliação externa (em busca de uma visão coerente do sistema tomado globalmente), disponibilizando aos professores instrumentos de classificação e de análise de resultados que providenciem informações úteis.^{34.35}

A perspetiva anteriormente explicitada tem na sua base preocupações de articulação entre a avaliação interna e a externa, tomando a última como uma inevitabilidade, porque consubstancia um instrumento de intervenção política que apresenta custos relativamente controlados, rapidez de implementação e que responde à vontade e controlo externos (Hamilton, 2003). Estas são, aliás, ideias também expressas na afirmação: “Unable to legislate classroom behavior, policy makers realize that testing is a tool that motivates educators to alter instructional process.” (Madaus, Russell & Higgins, 2009, p. 29).

Numa outra perspetiva, embora também ela considerando a relação entre estas duas modalidades de avaliação, poderá ser afirmado que a indução de práticas inovadoras de ensino, a moderação das avaliações internas e, ainda, a motivação acrescida que pode trazer às dinâmicas de ensino e de aprendizagem constituem efeitos de valor positivo que a investigação tem reconhecido à avaliação externa (Fernandes, 2014).

³⁴ A autora assenta a proposta no pressuposto de que os especialistas externos reúnem competências suplementares para os realizar melhor e que, mesmo havendo localmente capital de saber, esta abordagem será mais eficiente, porque menos consumidora de tempo.

³⁵ A abordagem referida, nomeadamente a linha de ação que propõe uma intervenção central no acompanhamento do sistema de avaliação interna, ganha alguma relevância no presente trabalho, considerando a natureza dos TI.

O ponto é que essa relação, positiva e sustentável, entre a avaliação interna e a externa terá de assentar na oportunidade de, através da avaliação externa, produzir informação consistente e divulgá-la de forma continuada, no tempo e nos meios, sem que da devolução de informação resultem apenas exercícios simplistas de comparações entre escolas ou regiões, ou entre os resultados da avaliação externa e os da interna. Neste âmbito particular, é necessário ter em conta que as relações entre uma e outra, nomeadamente no que respeita a resultados, não permitem comparações lineares – são, é bom lembrar, avaliações de natureza diferente, com propósitos diversos e formas de concretização também diferentes. Cada uma define, geralmente, objetos de avaliação que não são absolutamente coincidentes, sendo que, entre muitos outros aspetos, a avaliação interna é intrinsecamente, por regra em oposição à externa, uma avaliação de processo, contextualizada e diversificada nos instrumentos.

A verdade é que nem sempre a relação se revela *positiva e sustentável*. Em diversas situações, os objetivos serão verdadeiramente mais limitados: vários países desenvolvem e aplicam avaliações externas, em grande parte, devido à insatisfação quase crónica que se sente em relação ao sistema educativo. Utiliza-se, assim, a avaliação externa com o mero intuito de *regular/controlar* situações ou partindo das perceções, facilmente partilhadas pela sociedade (como já se referiu), de que a aplicação de testes em larga escala vai providenciar medidas fiáveis de desempenho dos alunos, documentar deficiências dos sistemas, vai provocar respostas positivas de alunos e professores e assegurar o desenvolvimento e a competitividade internacional (Madaus, Russell e Higgins, 2009).

A questão é que há uma natureza paradoxal na implementação de sistemas de avaliação em larga escala. E o paradoxo resulta, como explicitam os três autores que têm vindo a ser referidos³⁶, do facto de, perante a vontade de intervir positivamente no sistema – identificando, para ajudar, quem e o que não estará a desempenhar bem –, ser posto em marcha um conjunto de ações que visam melhorar os resultados, ações que simultaneamente têm impactos negativos (não procurados) nas escolas, alunos, professores e encarregados de educação.

³⁶Madaus, Russell & Higgins (2009).

Perante esta natureza paradoxal, o caminho será o de reconhecer e compreender a natureza e a extensão dos efeitos secundários, como se se tratasse de uma intervenção médica, que é poucas vezes cirúrgica na sua ação, mas que se presume e se deseja equilibrada nos seus prós e nos seus contras.

Ou é, se quisermos perspetivar de um outro ângulo, o chamado *efeito do observador*, em que o fenómeno ensino-aprendizagem-avaliação é alterado por via de estar a ser observado (avaliado), porque o que observamos é exposto a um método particular de questionamento, à utilização de um instrumento, e é, portanto, modificado. Nesta perspetiva, o caminho para mitigar o efeito será, entre outros aspetos, o de aperfeiçoar as técnicas e os instrumentos de observação.

Efetivamente, os programas de avaliação externa, enquanto intervenções de monitorização dos sistemas educativos, apresentam vários efeitos perversos, não desejados, e que não podem deixar de ser analisados à luz de um conjunto de impactos que representam para o sistema, a saber: na forma como as escolas e os professores desenvolvem o currículo, naquilo que é selecionado para ensinar e na forma como é ensinado, no modo como se trabalha com os diferentes alunos, bem como naquilo que se avalia e como. São os efeitos habitualmente identificados como *estreitamento do currículo* (e.g., Fernandes, 2005; Gipps, 2012), *curricular distortion* (Gipps, 2012) e, ainda, como *teaching to the test* (e.g., Stobart, 2008, Redfield, 2001, Gipps, 2012). A tudo isto acrescem os efeitos pessoais, sociais e políticos naturalmente associados à aplicação de instrumentos de avaliação externa de larga escala, com particular expressão para aqueles que têm funções de certificação e de seleção (as denominadas provas *high-stakes*, por oposição às chamadas *low-stakes* – sem implicações nos percursos escolares, com funções de controlo e de diagnóstico, mas, em algumas situações, igualmente merecedoras de uma projeção social e mediática assinalável).

Independentemente das razões que possam estar na base da implementação de programas de avaliação externa, será de não desconsiderar a assunção que apresentámos inicialmente: a de que o seu intuito é a melhoria das aprendizagens. Ora, o palco das aprendizagens é a escola, a sala de aula. E sobre esta relação, estaremos longe de poder simplesmente assumir que a avaliação externa (de boa qualidade) tem como consequência melhores aprendizagens na

sala de aula. Mehrens (2002) afirma precisarmos de evidências de que os usos e interpretações dos resultados da avaliação externa estão a contribuir para a melhoria das aprendizagens. Vai mais longe: questiona-se sobre a relação custo-benefício (entendendo aqui *custo* como custo financeiro dos programas). Dos vários aspetos que discute, destacamos três: (i) a capacidade de os resultados gerarem reformas curriculares e/ou alterações nas formas de ensinar; (ii) a relação entre melhores resultados e aprendizagem efetiva; (iii) os efeitos da divulgação alargada de resultados — geradora de credibilidade ou apenas informação para críticos?

Questão central para a abordagem da avaliação externa é a conceção dos instrumentos de avaliação e a sua aplicação. Deverá ser assegurado que os resultados gerados sejam vistos pelos vários utilizadores como merecedores de confiança. Existem dois critérios de qualidade que estarão na base dessa confiança: fiabilidade e validade. Seguimos Black & Wiliam (2006) na abordagem aos conceitos: partindo da premissa “No test is perfectly reliable.” (p. 120), afirma-se que é altamente improvável que a pontuação obtida por um aluno numa dada ocasião seja exatamente a mesma numa outra, ainda que no mesmo teste. No entanto, se o aluno realizasse o mesmo teste, ou teste semelhante, em várias ocasiões, então, a média das pontuações seria, genericamente, um bom indicador do que o teste estaria a medir. Porém, este bom indicador (*true score*) não significa que aquele aluno tenha, de facto, uma boa competência, por exemplo, de leitura (*true ability*). Equivale a explicação, de forma muito sintética, respetivamente, aos conceitos de fiabilidade (o rigor com que o teste mede o que pretende medir; a consistência de resultados) e de validade (em que medida o teste realmente avalia o que foi desenhado para medir).³⁷

³⁷ A questão abordada neste parágrafo também ocupa, naturalmente, lugar de relevo na análise de Madaus, Russell e Higgins (2009). A resposta à pergunta *O que é um teste?* é um contributo importantíssimo para a forma como se perspectivam, como se leem, os resultados de uma dada aplicação em larga escala. Um teste, designadamente um de *papel e lápis*, é limitado no seu objeto de avaliação, fará a representação de uma amostra de conhecimento ou de desempenho, e é usado para se fazerem inferências sobre o que alguém sabe ou sabe fazer sobre um dado domínio. Um teste providencia informação útil, mas falível.

Também Mehrens (1984) havia prescrito uma abordagem de prudência sobre leitura de resultados de testes: "The only reasonable, direct inference you can make from a test score is the degree to which a student knows the content that the test samples. Any inference about why the student knows that content to that degree...is clearly a weaker inference..." (Mehrens, 1984, p. 10).

Os programas de avaliação externa, apesar de se constituírem como uma intervenção, a vários níveis, complexa, especialmente no que respeita aos efeitos globais e, particularmente, na forma como se articulam com a avaliação interna, continuam a representar um peso significativo nos sistemas de ensino, porque as sociedades continuam a não conseguir encontrar alternativas credíveis, ou, se quisermos, continuam em busca de cenários que quantifiquem ou que meçam qualidade de forma homogénea ou equitativa. As limitações e as potencialidades dos sistemas de avaliação externa exigem da parte de quem tem responsabilidades em matéria de política educativa uma abordagem informada e prudente, para que se potencie a melhor apropriação e utilização dos resultados disponíveis. É fundamental ter presente que a sua existência não é garante de qualidade das aprendizagens e que os instrumentos que usam devem ser alinhados com o currículo, ser consistentes e apresentar qualidade psicométrica (Fernandes, 2014).

A coexistência da avaliação interna com a avaliação externa, a articulação entre uma e outra, permitirá tirar partido de uma ideia central aos processos de recolha de informação: a triangulação como forma de obter informações mais credíveis e com maior probabilidade de contribuírem para a melhoria da qualidade do que se ensina e do que se aprende.

CAPÍTULO 4 – METODOLOGIA

OPÇÕES METODOLÓGICAS E PROCEDIMENTOS PRINCIPAIS

O estudo assenta numa avaliação de cariz quantitativo, tendo como fonte primária três bases de dados de inscrição das escolas nos TI, abrangendo a oferta apresentada para o 3.º CEB e ES, relativa aos anos 2010/2011, 2011/2012 e 2012/2013. As Bases de Dados (BD) do Projeto TI do triénio apresentam um conjunto de informações administrativas, de participação e de caracterização sumária das escolas que aderiram ao Projeto.

A opção por centrar a análise na oferta de TI dirigida exclusivamente àqueles dois ciclos de escolaridade resultou do facto de, neste Projeto, estes serem os ciclos que partilham, no quadro conceptual anteriormente descrito, um conjunto de finalidades. De facto, no desenho do programa, as finalidades enunciadas, bem como os diversos procedimentos de implementação e de concretização do Projeto que as acompanham, tinham âmbito sobre a oferta proposta às escolas para o 3.º CEB e ES, não havendo coincidência absoluta com os propósitos, e mesmo com os procedimentos, dos TI do 1.º CEB. Desta forma, a opção visou dar solidez à análise, partindo de um conjunto de dados que, na sua origem, apresentam consistência por derivarem de uma base comum.

As bases de dados foram disponibilizadas pelo organismo gestor do Projeto TI – o IAVE, I.P., mediante protocolo estabelecido com o Instituto de Educação da Universidade de Lisboa (Anexo A)³⁸.

Nas referidas BD, para cada um dos anos, estão preenchidos diferentes campos que identificam a escola participante e o elenco de TI que a escola compôs para a sua participação. As informações convocadas para o presente estudo foram, no que respeita à escola: o nome e o código GEPE³⁹, o concelho, a natureza jurídica; e, relativamente à

³⁸ A fonte primária da quase totalidade dos dados analisados são as bases de dados disponibilizadas pelo IAVE, I.P., pelo que, ao longo do trabalho, apenas se identificam outras fontes usadas.

³⁹ Nomenclatura utilizada para a identificação dos códigos de escola no triénio entre 2010 e 2013. Atualmente, este código tem a designação de DGEEC – acrónimo da Direção Geral de Estatísticas da Educação e Ciência.

composição do elenco anual de TI, o(s) teste(s) intermédio(s) em que cada escola se inscreveu, bem como o período letivo de realização do teste selecionado. Esta base de dados contabiliza 37 660 registos de inscrições no triénio analisado.

As três bases anuais de inscrição das escolas nos TI foram, numa primeira fase de tratamento da informação, agregadas numa única base de dados, tendo a escola como unidade de registo. Nesta base de dados agregada, manteve-se o conjunto inicial de variáveis de caracterização das escolas (nome, código GEPE, concelho e natureza jurídica), acrescentando-se as variáveis relativas aos anos de participação no Projeto (um, dois ou três) e as variáveis respeitantes ao elenco de testes disponibilizados no conjunto dos três anos do Projeto TI. Esta segunda base de dados contabiliza 1378 registos – número de escolas que participaram pelo menos num dos anos do triénio analisado.

Numa segunda fase de tratamento da informação, foram selecionadas as escolas que, cumulativamente, escolheram realizar TI do 3.º CEB e do ES e que participaram em todos os anos do triénio em análise. Esta opção metodológica deveu-se à necessidade de garantir um conjunto de dados mais robustos para a análise de padrões de participação das escolas. Esta base de dados contabiliza 550 registos de escolas cujo denominador comum é reunirem condições para se inscreverem na totalidade de TI disponibilizados. A partir desta etapa, o estudo desenvolveu-se tendo por objeto este universo de escolas (apurado nas condições descritas).

No âmbito do quadro metodológico descrito, outras opções são explicitadas e discutidas, em forma de notas metodológicas, que acompanham, por facilidade de leitura, algumas fases da apresentação e análise de resultados constantes do capítulo seguinte.

Considerando as questões e a matriz de investigação definidas, foi traçada uma estratégia que pretendeu percorrer um plano para a análise dos dados disponíveis, cujas etapas sequencialmente permitissem: (i) apresentar e caracterizar o universo de escolas participantes no Projeto TI; e (ii) apresentar e caracterizar a adesão das escolas (definição da quantidade e variedade de TI selecionados; definição de padrões de participação).

Visando cumprir as etapas mencionadas e alcançar o objetivo proposto para a investigação, foram aplicados métodos de análise descritiva, recorrendo à representação gráfica, e ao método multivariado de análise fatorial, em particular à Análise em Componentes Principais –ACP)⁴⁰.

Em suma, tendo em consideração os propósitos da avaliação, foi definida a natureza dos dados a analisar, foram convocados os métodos que respondem de forma mais fiável e rigorosa ao objetivo de descrição, análise e interpretação da adesão das escolas ao Projeto. À informação gerada reconhece-se utilidade para a definição do destino do programa.

⁴⁰ A análise de dados foi suportada pelo *software* estatístico IBM SPSS Statistics (*Statistical Package for the Social Sciences*).

CAPÍTULO 5 – APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

A BASE DE DADOS DO PROJETO TESTES INTERMÉDIOS (TRIÊNIO 2010-2013)

A BD TI 2010-2013 apresenta 37 660 registos (considerando cada registo a opção de inscrição numa dada aplicação – TI), o que corresponde ao total de inscrições realizadas no conjunto de três anos letivos.

A referida BD apresenta registos por escola, enquanto unidade orgânica, a que corresponde, conforme codificação vigente de 2010 a 2013, um denominado código GEPE. A lógica de inscrição por escola manteve-se no período em análise, o que significa que, no caso dos agrupamentos de escolas, eram realizadas tantas inscrições quantas as escolas com 3.º ciclo do ensino básico e/ou secundário que os constituíam, sendo, desta forma, a escola a unidade de análise no presente trabalho.

Na primeira abordagem a esta fase de análise, a perspetiva é a da apresentação e caracterização das escolas participantes no Projeto, acompanhada de informação genérica sobre a sua adesão e representatividade no universo de escolas do país.

No presente trabalho, a adesão é o resultado da seleção que em cada ano letivo e no triénio as escolas realizaram perante o elenco de TI disponibilizado.

A configuração do elenco sofreu alterações ao longo dos três anos, conforme explicitado anteriormente, tendo-se mantido inalterado o princípio da adesão facultativa. As escolas voluntariamente realizavam as suas escolhas: (i) inscrevendo-se no Projeto (enquanto condição base de acesso ao programa TI) e (ii) escolhendo as disciplinas que a cada ano letivo pretendiam que fizesse parte do seu elenco individual de testes a aplicar⁴¹.

⁴¹ No triénio em análise, em algumas disciplinas, a escolha também pôde recair no número de testes a realizar, no mesmo ano, para uma mesma disciplina.

Ainda nesta fase de análise, uma segunda abordagem é perspectivada a partir da caracterização da adesão – centrada fundamentalmente no apuramento da escolha de TI pelas escolas participantes, considerando o ciclo de escolaridade e a disciplina⁴². Esta é uma fase em que se aborda o que poderão ser consideradas tendências de escolha, de que resulta um conjunto de testes, em representação de disciplinas, mais populares, que reuniram o consenso do universo de escolas participantes ao longo do triénio.

O UNIVERSO DE ESCOLAS PARTICIPANTES NO PROJETO TESTES INTERMÉDIOS

Considerando as escolas que participaram no Projeto, no triénio, tendo, portanto, em conta o total de anos de participação, verifica-se que a maioria das escolas esteve presente nas três edições (1215 unidades, correspondente a 88,2%) – Quadro 3.

Quadro 3 – Número de anos de participação das escolas no Projeto TI no triénio (2010-2013)

ANOS	N	%
1 ano	38	2,8
2 anos	125	9,0
3 anos	1215	88,2
Total de escolas que participaram pelo menos um ano no triénio 2011-2013	1378	100,0

O número de provas disponibilizadas foi sendo menor ao longo dos três anos. Conforme Gráfico 4 (e informação discriminada no Anexo C), a oferta, em 2011, foi de 21 provas, em 2012, de 18 provas, reduzindo-se, em 2013, para 14.

O primeiro ano do triénio (2010/2011) é o que apresenta um maior número de inscrições, estando este volume principalmente relacionado com o maior número de testes

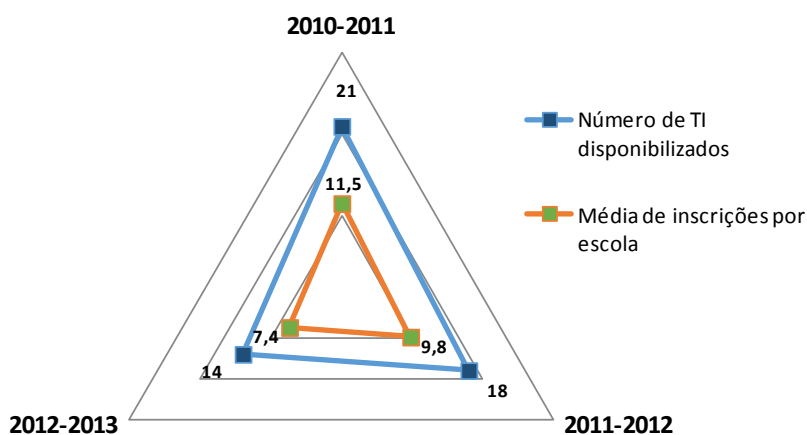
⁴² Sempre que do elenco anual de testes faça parte mais do que um teste por disciplina (correspondente a testes de diferentes anos de escolaridade) ou mais do que um teste por ano letivo (correspondente a, por exemplo, dois testes da mesma disciplina, no mesmo ano letivo), o tratamento será desagregado, indicando-se para cada situação o ano de escolaridade ou o ano de escolaridade e o teste. Seguir-se-á a codificação apresentada na Lista de abreviaturas e siglas, por exemplo, Mat A_11_t1 corresponde ao teste intermédio 1 (num ano letivo em que a oferta foi superior a um teste) da disciplina de Matemática A do 11.º ano de escolaridade.

disponibilizados nesse ano, comparativamente com os restantes anos em análise (Gráfico 4). Se se tiver em consideração o número de escolas participantes, verifica-se que é no segundo ano do triénio (2011/2012) que mais escolas aderiram ao Projeto TI, ano que coincide com a oferta mais diversificada⁴³.

No triénio analisado, no universo dos estabelecimentos de ensino básico e secundário, 1 378 escolas participaram em pelo menos um dos anos.

O número médio de inscrições por escola foi superior em 2010/2011 (11,5 testes, correspondente a um valor médio de adesão de 54,9%). Tendo em conta o total de provas disponibilizadas e a média de inscrições por escola, pode afirmar-se que a adesão foi expressiva em qualquer dos três anos e constante no triénio. Em média, as escolas participantes inscreveram-se em mais de metade do elenco disponibilizado e a variação entre os três anos em análise não ultrapassa 2pp.

Gráfico 4 – Número de provas disponibilizadas e média de inscrições por escola no triénio (2010-2013)



Ano letivo	Escolas participantes	Inscrições realizadas	Nº de provas disponibilizadas	Adesão média (em %)
2010-2011	1306	15045	21	54,9
2011-2012	1332	13000	18	54,2
2012-2013	1298	9615	14	52,9

⁴³ O ano letivo 2011/2012 é aquele que regista um elenco com o maior número de disciplinas (12 disciplinas).

Feita uma apresentação genérica do universo de escolas participantes no Projeto, avança-se, nesta fase, para a caracterização das escolas que aderiram no triénio 2010-2013.

A base de dados TI não apresenta informação sobre a tipologia de escola entendida como a tipificação do estabelecimento de ensino de acordo com o(s) ciclo(s) de escolaridade em cada ano oferecido(s) por cada unidade. Os dados disponíveis não permitem aquela tipificação.

Aliás, o Decreto-Lei n.º 299/2007, de 22 de agosto, na tipificação dos estabelecimentos de ensino, define as normas aplicáveis à denominação dos estabelecimentos de educação e ensino não superior públicos, estabelecendo, no seu artigo 8.º, que os estabelecimentos da rede pública são designados em função do nível de educação ou de ensino ou da modalidade que exclusiva ou prioritariamente ministram. Esta última disposição (*que [...] prioritariamente ministram* – que possibilita que a designação decorra também do ciclo de ensino ou da modalidade de formação que é prioritariamente ministrada no estabelecimento) implica, na prática, que, da designação da escola como secundária, não resulte informação conclusiva sobre a modalidade de ensino que oferece. De facto, uma escola de tipologia *escola secundária* pode oferecer 3.º CEB.

Complementarmente, as reconfigurações de que a rede escolar e as próprias unidades que a constituem foram objeto, nos anos em análise, e o facto de não se dispor de informação sobre a tipologia de escola (entendida como oferta de ciclos de escolaridade), à data da realização de cada elenco anual de TI, conduziu à opção de agrupar as escolas participantes de acordo com a sua participação, no caso concreto, de acordo com o(s) ciclo(s) de escolaridade dos testes que anualmente cada escola foi seleccionando para aplicação.

Trata-se, assim, de uma denominada *tipologia de participação*, em que as escolas são divididas da seguinte forma: Escolas TI ES, em que se incluem as que tenham seleccionado apenas TI do ensino secundário; Escolas TI EBS, em que se incluem as que tenham participado em TI do ensino básico e do ensino secundário; e Escolas TI EB, que agrupam as que tenham seleccionado exclusivamente TI do ensino básico.

Embora se reconheça as limitações de uma abordagem como a descrita, esta constitui-se como a que permite uma análise rigorosa dos dados disponíveis.

Nesta perspetiva, e de acordo com a explicitada *tipologia de participação*, o universo no triénio foi constituído maioritariamente (53,4%) por Escolas TI EB, a que se seguiram, as Escolas TI EBS – 42,4% e, com apenas 4,2%, as denominadas Escolas TI ES (Quadro 4).

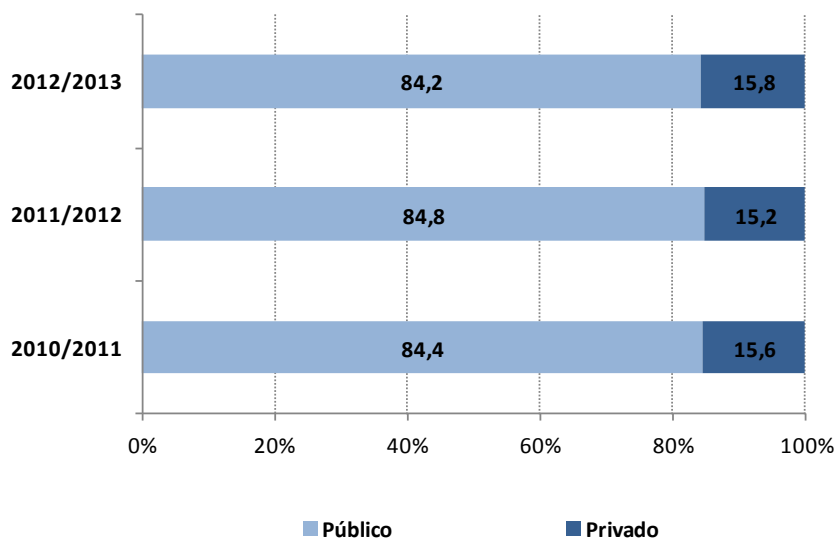
Da distribuição resultante do critério acima explicitado, não fica claro o significado da participação das Escolas TI EB e das TI ES. De facto, sem se poder apurar quantas escolas têm o seu âmbito de seleção de TI definido obrigatoriamente pelo(s) ciclo(s) de ensino que oferece(m), o significado da sua representação no universo em estudo é pouco relevante – esta é, em rigor, uma limitação que se impõe à leitura da participação das Escolas TI EB e TI ES. No entanto, situação bem distinta caracteriza 42,4% de escolas (correspondentes a 584 unidades) que apresentam inscrições em testes do 3.º CEB e em testes do ES. Para estas, a possibilidade de escolha é comprovadamente efetiva: são escolas que, oferecendo os dois ciclos de escolaridade, optaram por selecionar testes desses dois ciclos.

Quadro 4 – Escolas participantes por Tipologia de Participação no triénio (2010-2013)

Tipologia de participação	Triénio 2010-2013	
	N	%
Escola TI ES	58	4,2
Escola TI EBS	584	42,4
Escola TI EB	736	53,4
No conjunto dos três anos	1378	100,0

Se se considerar a natureza jurídica das escolas participantes (Gráfico 5), verifica-se que o grosso das escolas que aderiram ao Projeto TI é de natureza pública, aproximadamente 84,5%, sendo cerca de 15,5% a representação das escolas privadas. Esta relação mantém-se praticamente inalterada ao longo do triénio.

Gráfico 5 – Escolas participantes no Projeto TI, por ano e natureza jurídica (em %)



No que respeita à adesão das escolas participantes no Projeto TI desagregada por natureza jurídica (Quadro 5), verifica-se que os valores médios anuais são mais elevados nas escolas privadas, apresentando as escolas públicas uma adesão média que oscila entre 51,6% e 53,7%. Estes dados revelam diferenças que variam entre 7,7pp e 9pp nas médias anuais de adesão das escolas de uma e de outra natureza jurídica, com as escolas privadas a consistentemente selecionarem um maior número de testes intermédios para aplicação.

Registe-se ainda, à semelhança do que se havia verificado a partir dos dados do Gráfico 4 (sem desagregação das escolas por natureza jurídica), que a variação da adesão média entre os três anos é inexpressiva, o que revela fidelidade na participação das escolas, conclusão que também será legítima, considerando que a esmagadora maioria das escolas (1215 escolas – 88,2%) decidiu manter-se participante no Projeto TI durante os três anos.

Quadro 5 – Adesão média das escolas participantes no Projeto TI, por ano e natureza jurídica

Ano letivo	Natureza jurídica					
	Público			Privado		
	Escolas	Inscrições	Adesão média (em %)	Escolas	Inscrições	Adesão média (em %)
2010-2011	1102	12426	53,7	203	2619	61,4
2011-2012	1127	10771	53,1	202	2229	61,3
2012-2013	1090	7875	51,6	205	1740	60,6

As dificuldades anteriormente explicitadas, no que respeita à tipologia de escolas, estendem-se à definição da constituição da rede de escolas do ensino público e privado. Perante esta limitação, mas não querendo deixar de dar leitura à contribuição da participação das escolas públicas e das privadas no universo das escolas participantes no Projeto, optou-se por comparar o peso relativo de escolas públicas e privadas, em cada ano, com uma estimativa do peso relativo que cada uma destas naturezas jurídicas assume nas bases de dados anuais do Júri Nacional de Exames (JNE), relativas às provas finais e exames nacionais do 3.º CEB e do ES, respetivamente.

Assim, e para cada um dos anos que constituem o triénio em análise, fez-se o apuramento do número de escolas que realizaram a prova final da disciplina de Português (3.º CEB – 9.º ano) e o exame nacional de Português (ES – 12.º ano). A opção pela contabilização do número de escolas que realizaram aquelas provas de avaliação externa resulta do facto de estas apresentarem uma quase total cobertura, nestes ciclos. Desta forma, está a ser considerado que os registos constituem, para este efeito, o universo das escolas da rede escolar nacional, no qual se identificam as escolas do ensino público e do privado. Esta opção metodológica permite estimar a representatividade, no caso concreto, das escolas participantes por natureza jurídica, embora do apuramento resultem valores por estimativa, uma vez que da contabilização de escolas com provas finais de 9.º ano (apuramento para a PF de Português na base ENEB) e de escolas com exames nacionais do 12.º ano (apuramento para o Ex de Português na base ENES) resultam dois universos de escolas que sabemos não serem exclusivos, isto é, entre os grupos existem sobreposições, considerando que há escolas que realizam as provas de 9.º e de 12.º ano. Esta seria uma limitação facilmente ultrapassável se as duas bases fossem direta e facilmente consolidáveis, por exemplo se registassem campos equivalentes/iguais para entradas iguais ou se, pelo menos, registassem ambas de forma nominalmente coerente as entradas.

O Quadro 6 apresenta, por ano letivo, o número de escolas por natureza jurídica – pública e privada – que realizaram a PF de Português de 9.º ano e o número das que realizaram o Ex de 12.º ano de Português. Acresce a esta informação a representatividade de cada natureza

jurídica no universo de escolas que em cada ano realizaram aquelas provas de avaliação externa.

Neste mesmo Quadro, verifica-se que as escolas privadas, no caso das escolas que realizaram provas de avaliação externa, considerados os três anos, representam 1/5 do universo de escolas da rede escolar nacional, distribuição que o universo de escolas participantes no Projeto TI acompanha, uma vez que, no triénio em análise, conforme referido anteriormente, participaram cerca de 84,5% de escolas públicas e 15,5% de privadas.

Quadro 6 – Distribuição de escolas que realizaram Provas Finais do 9.º ano e Exames Nacionais do 12.º ano no triénio 2010-2013, por natureza jurídica

Ano letivo	Público				Privado				Total	
	9.º ano Provas Finais		12.º ano Exames		9.º ano Provas Finais		12.º ano Exames		Total	
	N	%	N	%	N	%	N	%	3.º CEB	ES
2010/2011	1080	83,7	493	80,2	211	16,3	122	19,8	1291	615
2011/2012	1114	83,9	500	80,3	214	16,1	123	19,7	1328	623
2012/2013	1086	83,0	501	80,2	222	17,0	124	19,8	1308	625

Fonte: JNE, 2011, 2012 e 2013

O Quadro 7 apresenta a distribuição das escolas participantes no Projeto TI por NUTS III, identificando ainda as Escolas Portuguesas no Estrangeiro (EPE), e as médias de inscrição por escola, no triénio 2010-2013.

Os dados apresentados no quadro permitem verificar que as duas grandes áreas metropolitanas de Lisboa e Porto e as suas regiões limítrofes (respetivamente, Península de Setúbal, Ave e Tâmega), além do Algarve, apresentam o mais elevado número de escolas participantes no Projeto TI⁴⁴. A contribuição da sub-região da Grande Lisboa e do Grande Porto para o universo de escolas participantes é, em média, 16,8% e 10,8%, respetivamente. A distribuição apresentada no Quadro 7, mostra, ainda, que, embora com diferentes contribuições, todas as regiões tiveram escolas representadas no Projeto, no triénio 2010-2013, o que constitui indicador da sua expressão nacional.

⁴⁴ O que não surpreende atendendo à densidade populacional e à correspondente oferta de escolas da rede pública e privada.

Quadro 7 – Média de inscrições por NUTS III no triénio 2010-2013

NUTS III	Inscrições	Escolas inscritas	% de escolas inscritas	Número médio de inscrições por escola
Região Autónoma dos Açores	896	30	2,2	29,9
Região Autónoma da Madeira	666	33	2,4	20,2
Minho-Lima	872	33	2,4	26,4
Cávado	1110	49	3,6	22,7
Ave	1514	66	4,8	22,9
Grande Porto	3989	149	10,8	26,8
Tâmega	1897	66	4,8	28,7
Entre Douro e Vouga	921	32	2,3	28,8
Douro	918	31	2,2	29,6
Alto Trás-os-Montes	861	29	2,1	29,7
Baixo Vouga	1596	59	4,3	27,1
Baixo Mondego	1406	49	3,6	28,7
Pinhal Litoral	1072	35	2,5	30,6
Pinhal Interior Norte	638	22	1,6	29,0
Pinhal Interior Sul	301	7	0,5	43,0
Dão-Lafões	1343	42	3,0	32,0
Serra da Estrela	215	8	0,6	26,9
Beira Interior Norte	610	20	1,5	30,5
Beira Interior Sul	340	11	0,8	30,9
Cova da Beira	466	14	1,0	33,3
Oeste	1225	41	3,0	29,9
Médio Tejo	950	28	2,0	33,9
Grande Lisboa	6449	231	16,8	27,9
Península de Setúbal	2267	87	6,3	26,1
Lezíria do Tejo	846	32	2,3	26,4
Alentejo Litoral	500	19	1,4	26,3
Alentejo Central	577	25	1,8	23,1
Alto Alentejo	715	27	2,0	26,5
Baixo Alentejo	699	26	1,9	26,9
Algarve	1490	69	5,0	21,6
EPE	311	8	0,6	38,9
Total	37660	1378	100,0	27,3

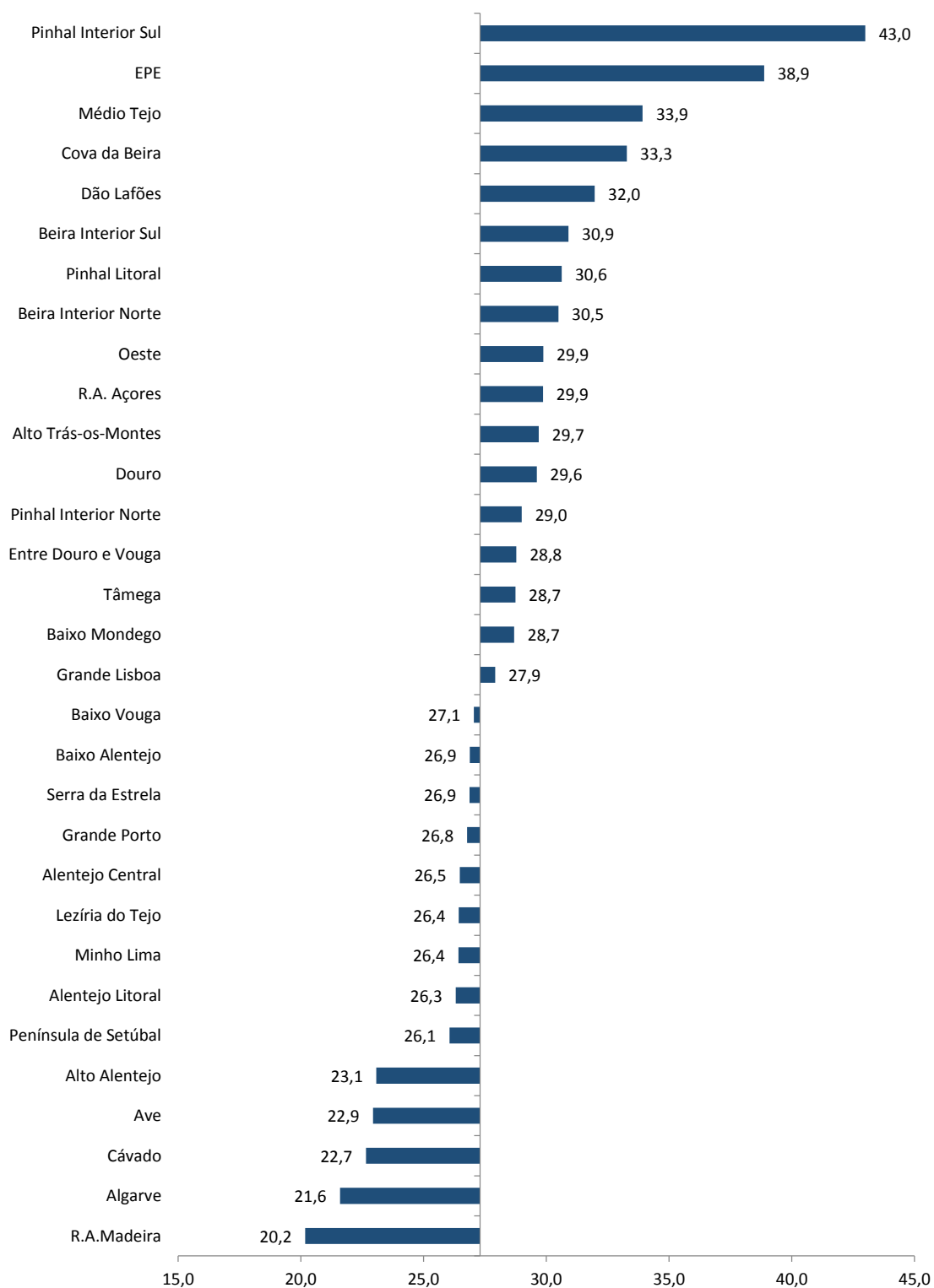
Se a distribuição for analisada em função do número médio de inscrições das escolas pertencentes a cada uma das sub-regiões e EPE (Gráfico 6), os dados revelam que são a sub-região do Pinhal Interior Sul, as EPE, e as sub-regiões do Médio Tejo e da Cova da Beira que apresentam as escolas que, em média, mais testes escolheram do elenco disponível. Sabendo que cada uma das regiões apresenta escolas das diversas tipologias (EB, ES e EBS), o que significa que nem todas as escolas das regiões estavam habilitadas escolher testes dos dois ciclos de escolaridade, e não esquecendo que a análise assenta num valor médio, não

deixam de ser significativo os valores de 81,1%, 73,4%, 64,0% e 62,8%⁴⁵, que representam em percentagem o número médio de provas realizadas, respetivamente, por cada uma das regiões (e EPE) referidas. São valores que indicam uma escolha quantitativamente expressiva: as escolas destas regiões, bem como o conjunto formado pelas escolas portuguesas no estrangeiro, em média, escolheram para o seu elenco individual muitos testes.

Na ordenação da média de inscrições no Projeto, apresentada por este mesmo gráfico (Gráfico 6), fica igualmente evidente uma outra distribuição: das 31 unidades consideradas, a maioria (17) regista médias de inscrição acima da média global (27,3), estando as restantes abaixo deste valor, com destaque para a Região Autónoma da Madeira, que apresenta, no triénio, a média mais baixa de inscrições por escola: 20,2.

⁴⁵ Correspondentes a 43; 38,9; 33,9 e 33,3, respetivamente (Gráfico 6).

Gráfico 6 – Ordenação da média de inscrições no Projeto TI por NUTS III (2010-2013)



Nota: A média global de inscrições por escola no triénio considerado foi de 27,3 (as barras à esquerda representam as NUTS III que registam uma média de inscrições abaixo da média global).

A ADESÃO DAS ESCOLAS AO PROJETO

Perspetivando agora o comportamento do universo de escolas participantes a partir da caracterização da adesão, isto é, centrando a análise na construção da resposta à questão *Que testes escolheram as escolas no triénio 2010-2013?*, será importante, antes de mais, recuperar, opções metodológicas tomadas, para as continuar a explicitar.

Recorde-se, nesta fase, que o apuramento que permitiu definir um universo de escolas que apresentam condições para aderir aos TI de cada um dos dois ciclos de escolaridade teve como objetivo primeiro estimar a representatividade da natureza jurídica das escolas. Atingido esse objetivo, prosseguiu-se com uma segunda fase de verificação para que se pudesse definir um universo de escolas que apresentassem condições para se inscreverem não só nas provas de Português (9.º e 12.º anos), mas também em cada uma das outras disciplinas em que existiu oferta TI, em cada ano (2010-2013). Para este efeito, foi verificada a existência em cada uma das escolas de exames relativos a cada uma das disciplinas, além de Português (exame com o código 639), que apresentam oferta TI no ES (Biologia e Geologia, Física e Química A e Matemática A). No caso da disciplina de Filosofia, uma vez que é uma disciplina comum à componente geral dos cursos científico-humanísticos, considerou-se que a sua oferta na rede escolar está, à partida, assegurada.

A verificação daquela condição só é relevante no caso do ensino secundário (é neste ciclo que as alternativas de oferta das disciplinas da formação específica existem), sendo que no caso do 3.º CEB a oferta curricular é globalmente uniforme, cobrindo necessariamente as disciplinas que faziam parte do elenco TI no triénio – além de Português (9.º ano), Matemática, Físico-Química, Ciências Naturais, Inglês, Geografia e História.

Já anteriormente se apresentou o número de escolas que se inscreveram em TI do 3.º CEB e em TI do ES, em cada um dos três anos. A variação do número de inscrições é naturalmente influenciada pelo número total de testes disponíveis em cada ano (21, 18 e 14 testes, respetivamente, no triénio), variando o número de escolas que aderiram aos testes de cada

um dos ciclos de escolaridade de forma pouco expressiva (entre 1233 e 1272, para o 3.º CEB, e entre 601 e 621, para o ES).

Na sequência do que foi anteriormente explicitado, o cálculo da percentagem de escolas que aderiram aos testes intermédios teve como referência o conjunto de estabelecimentos de ensino onde, no ano letivo correspondente, se realizou a Prova Final (PF) de Português do 9.º ano de escolaridade e o Exame Nacional (Ex) de Português do ensino secundário (neste caso, foram considerados os alunos internos e externos inscritos na primeira fase).

Deste exercício, que permite apurar objetivamente, para cada ano, o número de escolas que eram elegíveis como escolas participantes, resultam taxas de cobertura, tanto no ensino básico como no secundário, sempre iguais ou superiores a 94,8% (destaque-se o valor de 99,4%, para a cobertura do ensino secundário, em 2012/2013), o que é revelador da adesão maciça das escolas ao Projeto. Este é um comportamento que se mantém constante ao longo do triénio, como já tinha sido verificado na análise inicial (Quadro 8).

Quadro 8 – Adesão das escolas ao Projeto TI, por ano e ciclo de escolaridade

Ano letivo	Ciclos de escolaridade					
	Ensino Básico			Ensino Secundário		
	Inscrições	Escolas	% de cobertura*	Inscrições	Escolas	% de cobertura*
2010-2011	8515	1233	95,5	6522	601	97,7
2011-2012	7384	1272	95,8	5603	618	99,2
2012-2013	5771	1240	94,8	3921	621	99,4

Nota: *O cálculo da percentagem de escolas que aderiram aos TI teve como referência o conjunto de estabelecimentos de ensino onde, no ano letivo correspondente ao do teste em análise, se realizou a PF de Português do 9.º ano de escolaridade e o Ex de Português do ensino secundário (foram considerados os alunos internos e externos inscritos na primeira fase). A opção deve-se à larga cobertura que as duas provas de Português assumem na avaliação externa nacional.

Os três quadros que, de seguida, se apresentam (Quadros 9, 10 e 11) mostram a distribuição anual do número de escolas por teste intermédio. Neste momento, a análise centra-se na escolha de testes, ordenando-os de acordo com aquilo que se poderia considerar a sua popularidade, medida pela adesão das escolas. A taxa de escolas cobertas pelo Projeto TI segue a opção metodológica já explicitada anteriormente. O elenco anual de TI encontra-se dividido entre 3.º CEB e ES.

No primeiro ano do triénio, o elenco de provas repartiu-se por 12 testes do ensino secundário e nove do ensino básico. No conjunto de provas dirigidas ao ensino secundário, a esmagadora maioria das disciplinas apresenta taxas de cobertura superiores a 90%. As disciplinas de Biologia e Geologia, Física e Química A e Matemática A foram as que registaram, nos três anos, um número mais elevado de inscrições. Note-se que, no elenco de testes do ensino secundário, o número de inscrições por disciplina, coincidente com o número de escolas que em cada ano escolheu aderir à realização de um dado TI, não apresenta oscilações significativas entre as disciplinas (a diferença da taxa de cobertura não ultrapassa, na esmagadora maioria das disciplinas, os 5pp), o que, à partida, não permite outras leituras que não a de que existiu uma adesão expressiva a todos os TI, independentemente do ano de escolaridade (10.º, 11.º ou 12.º anos), isto é, a adesão não apresenta dependência do facto de haver coincidência ou não com o ano de realização de exame nacional.

Exceção a esse padrão é a disciplina de Filosofia (TI de 10.º ano), que apresenta, em 2010/2011, apenas 352 inscrições, o que equivale a uma cobertura de 57,2%. Nesta disciplina, 2010/2011 foi o primeiro ano de oferta de teste intermédio, sendo que nesse mesmo ano foi incluído no elenco de exames nacionais do ensino secundário o exame de Filosofia (11.º ano), com carácter opcional⁴⁶.

No ensino básico, sobressai a disciplina de Matemática (do 8.º e do 9.º ano), logo seguida pela de Português (9.º ano), ambas disciplinas com provas de avaliação externa. Este é, aliás, um padrão que se mantém ao longo dos três anos. É sobretudo nas disciplinas que apresentam provas finais, no caso do ensino básico, ou exames nacionais obrigatórios, no ensino secundário, que se regista um maior número de inscrições (Quadros 9, 10 e 11). Note-se que este comportamento se mantém, mesmo na situação em que o teste intermédio não coincide com o ano de prova de avaliação externa, como acontece com o teste de Matemática do 8.º ano, que apresenta um número de inscrições próximo do do teste equivalente do 9.º ano, ambos com taxas de cobertura de aproximadamente 90%.

⁴⁶ Conforme disposto no Decreto-Lei n.º 50/2011, de 8 de abril, e na Portaria n.º 244/2011, de 21 de junho de 2011.

Os TI das restantes disciplinas que compõem o elenco de oferta do ensino básico no triénio em análise (Ciências Naturais, Físico-Química, Geografia, História e Inglês) apresentam taxas de cobertura nunca inferiores a 50%, oscilando entre 64,9% (Físico-Química) e 52,7% (História) Registe-se, igualmente, que na ordenação do número de inscrição nestes TI (correspondentes a disciplinas que não apresentam provas finais) fica evidente a existência de um outro padrão: chamadas a seleccionar a partir do elenco disponível, as escolas, nos três anos, deram prioridade aos TI da disciplina de Físico-Química, logo seguida pela disciplina de Ciências Naturais. Neste *ranking* de inscrições, os TI das disciplinas de Geografia e História apresentaram, no triénio, os níveis mais baixos de adesão.

Quadro 9 – Número ordenado de inscrições, por Teste Intermédio e ciclo de escolaridade (2010/2011)

Testes Intermédios	N.º de inscrições	% de cobertura
TI – ES		
Biologia e Geologia – 11.º Ano_t1	573	93,5
Biologia e Geologia – 11.º Ano_t2	573	93,5
Física e Química A – 11.º Ano_t1	567	92,2
Matemática A – 12.º Ano_t2	566	93,4
Física e Química A – 11.º Ano_t2	565	91,9
Matemática A – 12.º Ano_t1	563	92,9
Matemática A – 10.º Ano	560	92,4
Matemática A – 11.º Ano_t1	556	91,7
Matemática A – 11.º Ano_t2	554	91,4
Física e Química A – 10.º Ano	548	89,1
Biologia e Geologia – 10.º Ano	545	88,9
Filosofia – 10.º Ano	352	57,2
TI – 3.º CEB		
Matemática – 9.º Ano_t1	1197	92,7
Matemática – 9.º Ano_t2	1189	92,1
Português – 9.º Ano	1165	90,2
Matemática – 8.º Ano	1150	89,1
Físico-Química – 9.º Ano	838	64,9
Ciências Naturais – 9.º Ano	796	61,7
Inglês – 9.º Ano	794	61,5
Geografia – 9.º Ano	711	55,1
História – 9.º Ano	680	52,7

No Quadro 10 ordena-se, por número de inscrições, os TI pertencentes ao elenco do ano letivo 2011/2012.

Quadro 10 – Número ordenado de inscrições, por Teste Intermédio e ciclo de escolaridade (2011/2012)

Testes Intermédios	Nº de inscrições	% de cobertura
TI – ES		
Biologia e Geologia - 11.º Ano	591	95,5
Física e Química A - 11.º Ano	588	95,1
Matemática A - 12.º Ano_t1	584	95,0
Matemática A - 12.º Ano_t2	583	94,8
Matemática A - 10.º Ano	571	92,8
Física e Química A - 10.º Ano	564	91,3
Matemática A - 11.º Ano	564	91,7
Português - 12.º Ano	561	90,0
Biologia e Geologia - 10.º Ano	554	89,5
Filosofia - 11.º Ano	443	71,1
TI – 3.º CEB		
Matemática - 9.º Ano	1246	93,8
Português - 9.º Ano	1204	90,7
Matemática - 8.º Ano	1198	90,2
Físico-Química - 9.º Ano	825	62,1
Ciências Naturais - 9.º Ano	819	61,7
Inglês - 9.º Ano	742	55,9
Geografia - 9.º Ano	681	51,3
História - 9.º Ano	676	50,9

No Quadro 11, lista-se, por ordem de número de inscrições, os TI do terceiro ano do triénio em análise (2012/2013).

Quadro 11 – Número ordenado de inscrições por Teste Intermédio e ciclo de escolaridade (2012/2013)

Testes Intermédios	Nº de inscrições	% de cobertura
TI – ES		
Física e Química A - 11.º Ano	592	94,6
Biologia e Geologia - 11.º Ano	591	94,1
Matemática A - 12.º Ano_t2	584	94,7
Matemática A - 12.º Ano_t1	579	93,8
Matemática A - 11.º Ano	568	92,1
Português - 12.º Ano	559	89,4
Filosofia - 11.º Ano	417	66,7
TI – E. Básico		
Matemática - 9.º Ano	1225	93,7
Português - 9.º Ano	1177	90,0
Físico-Química - 9.º Ano	732	56,0
Ciências Naturais - 9.º Ano	711	54,4
Inglês - 9.º Ano	689	52,7
História - 9.º Ano	601	45,9
Geografia - 9.º Ano	590	45,1

O UNIVERSO DE ESCOLAS DE TIPOLOGIA DE PARTICIPAÇÃO TI EBS

Uma análise mais detalhada, apontando para a definição de padrões de participação das escolas no Projeto exigiu, no quadro das opções metodológicas tomadas, que do universo de escola participantes (cuja apresentação e caracterização foi genericamente feita na secção anterior), se evoluísse para um conjunto de escolas que apresentavam, no triénio, uma tipologia de participação equivalente e, portanto, analisável na sua constituição e evolução, com o rigor pretendido.

Visando, então, uma maior consistência do estudo apresentado, extraiu-se do universo de escolas participantes o conjunto de escolas que cumulativamente satisfaziam as seguintes condições: (i) terem participado nos três anos; e (ii) serem escolas TI EBS, conforme classificação de acordo com a denominada tipologia de participação (explicitação feita anteriormente).

Considerando os dois critérios cumulativamente, do conjunto de 1378 escolas apuradas, destacaram-se 550 escolas com tipologia mista (EBS) e um histórico de três anos de participação contínua no triénio 2010-2013.

A partir desta fase, o trabalho centrou-se neste universo, que corresponde a aproximadamente 45% da participação total (1215 escolas participaram continuamente nos três anos). Numa primeira etapa, procedeu-se à caracterização das EBS por natureza jurídica e localização geográfica, seguindo procedimentos idênticos aos já operacionalizados para o conjunto total de escolas participantes. Avançou-se, de seguida, para uma análise centrada na análise da quantidade e variedade de TI selecionados pelas EBS, bem como na relação entre estas duas variáveis. Numa segunda etapa, foi então possível evoluir para um exercício de identificação de padrões de adesão, a partir da análise fatorial de componentes principais.

CARACTERIZAÇÃO DAS ESCOLAS TI EBS

No Quadro 12, em que se desagrega a natureza jurídica das TI EBS e se apresenta as médias de inscrições correspondentes, verifica-se que as TI EBS públicas representam 80,2% e as privadas 19,8%, constituindo uma distribuição não muito diferente da apurada para todas as escolas participantes, embora com um peso acrescido de 5pp para as escolas privadas.

As inscrições das TI EBS representam 62,4% do total de todas as inscrições no triénio, assumindo esta tipologia de escolas um relevo expressivo no conjunto de escolas participantes no Projeto. A média de inscrições e a adesão média são mais elevadas nas TI EBS (80,7%, de total de adesão média) do que no total de escolas analisadas por ano (valores a oscilarem entre 51,6% e 61,4%, conforme apresentado em secção anterior). Embora as tendências se mantenham, as escolas privadas registam uma média de inscrições por escola e uma taxa de adesão média ligeiramente superiores às encontradas para as escolas públicas, (Quadro 12).

Quadro 12 – Média de inscrições das Escolas TI EBS por natureza jurídica

Natureza jurídica	Escolas TI EBS			
	Escolas	Inscrições	Média de inscrições	Adesão média (%)
Público	441	18813	42,7	80,5
Privado	109	4705	43,2	81,4
Total	550	23518	42,8	80,7

A análise da participação no Projeto das TI EBS, no triénio, pelas NUTS III (e EPE) (Quadro 13) revela uma distribuição que cobre o território nacional, com as regiões da Grande Lisboa e Grande Porto, Península de Setúbal e Tâmega a apresentarem o maior número de escolas inscritas, numa distribuição semelhante à apurada para o conjunto total de escolas participantes.

No que respeita ao número médio de inscrições por escola e à correspondente taxa de adesão, os valores acima de 85% (que correspondem a um número médio de inscrições por escola superior a 45,1 – para uma oferta total no triénio de 53 testes) registam-se nas sub-regiões da Beira Interior Sul, Pinhal Interior Sul, Cova da Beira, Médio Tejo, Serra da Estrela,

Dão-Lafões e Baixo Mondego, apresentando cinco destas seis unidades uma concentração assinalável na região Centro (NUTS II).

A adesão média das escolas em cada sub-região (NUTS III) apresenta globalmente valores muito elevados, mesmo considerando as sub-regiões que, ao contrário das anteriormente referidas, se encontram no extremo oposto de uma ordenação que tem em conta o número médio de inscrições por escola. A RAA, a sub-região do Alentejo Litoral, a RAM e a sub-região do Douro registam taxas abaixo de 75%, pertencendo o valor menos elevado a esta última, com 70%.

Quadro 13 – Média de inscrições das Escolas TI EBS que participaram nos três anos do Projeto TI, por NUTS III

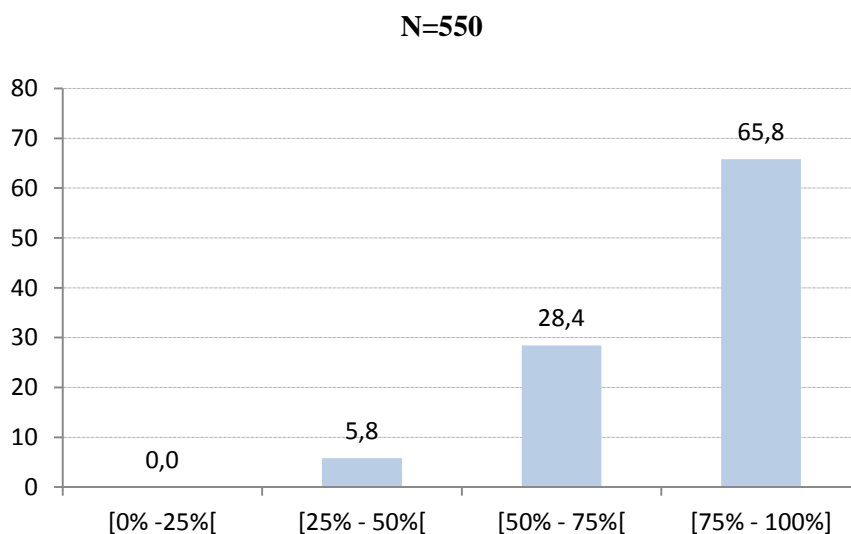
NUTS III	Inscrições	Escolas inscritas	Número médio de inscrições por escola	Adesão média (%)
Região Autónoma dos Açores	429	11	39,0	73,6
Região Autónoma da Madeira	791	21	37,7	71,1
Minho-Lima	600	15	40,0	75,5
Cávado	590	15	39,3	74,2
Ave	692	16	43,3	81,6
Grande Porto	2667	66	40,4	76,2
Tâmega	1242	28	44,4	83,7
Entre Douro e Vouga	582	13	44,8	84,5
Douro	705	19	37,1	70,0
Alto Trás-os-Montes	689	17	40,5	76,5
Baixo Vouga	996	24	41,5	78,3
Baixo Mondego	831	18	46,2	87,1
Pinhal Litoral	721	16	45,1	85,0
Pinhal Interior Norte	395	9	43,9	82,8
Pinhal Interior Sul	344	7	49,1	92,7
Dão-Lafões	975	21	46,4	87,6
Serra da Estrela	93	2	46,5	87,7
Beira Interior Norte	401	9	44,6	84,1
Beira Interior Sul	202	4	50,5	95,3
Cova da Beira	333	7	47,6	89,8
Oeste	718	16	44,9	84,7
Médio Tejo	610	13	46,9	88,5
Grande Lisboa	3968	90	44,1	83,2
Península de Setúbal	1482	34	43,6	82,2
Lezíria do Tejo	488	11	44,4	83,7
Alentejo Litoral	232	6	38,7	73,0
Alentejo Central	204	5	40,8	77,0
Alto Alentejo	418	10	41,8	78,9
Baixo Alentejo	451	11	41,0	77,4
Algarve	382	9	42,4	80,1
EPE	287	7	41,0	77,4
Total	23518	550	42,8	80,7

Focando a análise no número de TI que cada escola escolheu realizar no triénio, verifica-se que a maioria das escolas TI EBS (65,8%) realizou 3/4 ou mais do total de testes disponíveis (Gráfico 7). Deste grupo de escolas, 97 (17,6%) participaram em todos os testes disponibilizados (53 provas ao longo dos três anos em análise).

Da informação apresentada pelo Gráfico 7, importa ainda destacar o facto de não se registarem escolas cuja seleção de TI corresponda a menos 25% do elenco oferecido, bem como ser residual (5,8%) o número das que decidiram optar por realizar 26 ou menos dos 53 testes do triénio.⁴⁷

Os dados apresentados configuram um comportamento de seleção significativa: as EBS aderem ao Projeto e compõem o seu elenco com um número expressivo de TI.

Gráfico 7 – Testes realizados nas Escolas TI EBS no triénio 2010-2013 (em %).



Também no que respeita ao elenco de disciplinas, as EBS participaram de forma alargada, com mais de 71% das escolas a realizar pelo menos uma prova de todas as disciplinas

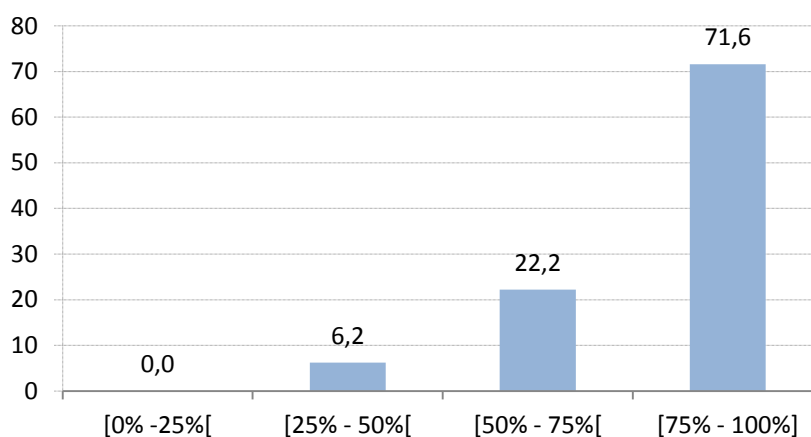
⁴⁷ Uma análise desagregada por natureza jurídica mostra uma distribuição por intervalos de adesão idêntica à dos resultados tomados globalmente. Não obstante o exposto, o intervalo [25% – 50%[tem uma maior representação das escolas privadas (11,0%, contra 4,5% das públicas). Nos outros dois intervalos os valores de adesão a número de TI não apresentam diferenças expressivas entre escolas públicas e privadas (no intervalo que agrega a escolha de entre 50% a 75% dos TI, há 29,3% de escolas públicas e 24,8% de privadas, sendo o intervalo superior – escolas que escolheram mais de 75% dos testes – ocupado por 66,2% de públicas e 64,2 de privadas).

disponibilizadas no conjunto dos três anos (Gráfico 8). Neste grupo, um número muito expressivo de escolas, 58% (319 Escolas TI EBS), não prescindiu de participar na avaliação de pelo menos uma prova de todas as disciplinas que tiveram TI no triénio.

A composição do elenco TI por cada escola, que agrega escolhas de quantidade e de variedade (o número de testes e as disciplinas cobertas pelos testes), revela uma clara tendência: a opção por um leque muito alargado e diversificado.

Gráfico 8 – Disciplinas cobertas nas Escolas TI EBS no triénio 2010-2013 (em %).

N=550



A relação entre as variáveis *taxa de adesão ao conjunto de TI* e *taxa de adesão ao leque de disciplinas disponíveis* está representada no Gráfico 9. Os pontos correspondem às frequências observadas para cada escola participante – cada par de valores representa o número de testes realizados e o número de disciplinas selecionadas.

Existe uma correlação forte, significativa e no mesmo sentido entre o número de testes realizados ao longo dos três anos e o número total de disciplinas selecionadas (R de Pearson = 0,784; $P < 0,001$). As escolas que realizaram um mais elevado número de testes também se inscreveram num leque mais variado de disciplinas.

Esta correlação, apesar de natural e expectável, não é a única possível. Para que esta fosse a única possível teríamos de estar perante um elenco que apresentasse uma distribuição

equitativa do número de testes por disciplina. Este não é o caso. Como se pôde verificar anteriormente, há disciplinas em que se disponibilizaram um maior número de testes do que noutras.

Registe-se, com base no que foi anteriormente explicitado a propósito da composição do elenco TI 2010-2013, um conjunto de leituras sobre os dados, agora perspetivados em função da relação entre *taxa de adesão ao conjunto de TI* e *taxa de adesão ao leque de disciplinas disponíveis*:

- com uma oferta de 53 testes e de 12 disciplinas, se a relação entre testes e disciplinas fosse proporcional, cada uma das disciplinas teria 4,4 testes;
- há quatro disciplinas (Matemática 3.º CEB, Biologia e Geologia, Física e Química A e Matemática A) que ofereceram mais do que um teste por ano, concentrando estas quatro disciplinas 30 testes (dos 30, 12 são testes de Matemática A);
- existindo 30 testes a cobrirem quatro disciplinas, este é o número máximo de testes que uma escola pode fazer com um número mínimo de disciplinas;
- a inscrição em 35 testes (11+12+12 – 66% da oferta) é suficiente para cobrir todas as disciplinas.

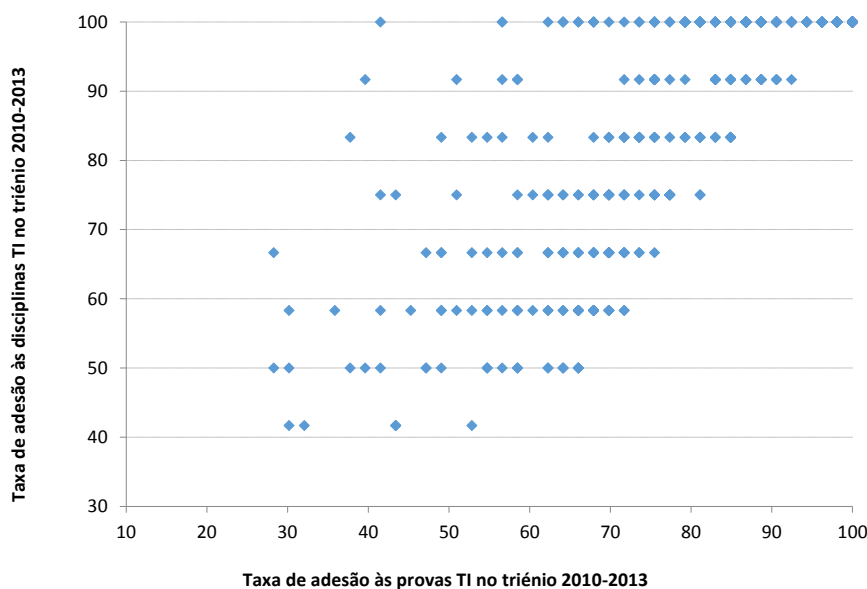
A análise do elenco revela quatro padrões de conjugação de disciplinas e de número de testes, com cada um dos padrões a revelar relações distintas entre as duas variáveis. Há um padrão que corresponde a uma disciplina com dois testes (A); um segundo padrão em que se encontram as disciplinas com três testes (B); um terceiro com cada disciplina a apresentar seis testes (C); e, por fim, um último padrão com uma disciplina a disponibilizar 12 testes (D)⁴⁸

⁴⁸ Veja-se uma distribuição dos referidos padrões na tabela que se apresenta:

Padrões	Disciplinas	N.º de testes por disciplina	Total de testes por padrão
A	Port_12	2	2
B	CFQ_9 / CN_9 / Geo_9 / Hist_9 / Ing_9 / LP_9 Filos (10.º e 11.º anos)	3	21
C	Mat (8.º e 9.º anos) BG (10.º e 11.º anos) / FQA (10.º e 11.º anos)	6	18
D	Mat A (10.º, 11.º e 12.º anos)	12	12

Efetivamente, a escolha de um mais elevado número de testes não implica obrigatoriamente que dessa escolha resulte uma maior variedade de disciplinas, como se pode observar no exemplo de Matemática A, disciplina em que é dada a possibilidade de as escolas se inscreverem em 12 testes (correspondentes a 22,6% do número de testes disponíveis) sem que haja mais do que uma disciplina coberta.

Gráfico 9 – Relação entre a quantidade de TI e a variedade de TI no triénio (2010-2013)



PADRÕES DE ADESÃO

A decisão da direção das escolas em participar no Projeto TI tem subjacente a vontade de selecionar um elenco de testes, como já ficou claro. O conjunto de informação até agora reunido confirma uma adesão muito expressiva e muito fidelizada das escolas a este Projeto. A finalidade com que aderem não será certamente comum a todas as escolas, nem mesmo se terá mantido inalterada ao longo dos três anos em análise.

Para algumas escolas, a motivação para a adesão estará relacionada com a preparação de alunos para as provas de avaliação externa, para outras, com a possibilidade de ter elementos que lhes permitam comparar a avaliação interna das aprendizagens com resultados gerados por instrumentos de conceção externa e de aplicação nacional, para outras ainda, a mobilização resultará mais da hipótese de, ao terem acesso a instrumentos de avaliação concebidos por uma estrutura externa credenciada, isso poder permitir a cada professor aferir o desempenho dos seus alunos por referência a padrões de âmbito nacional e/ou ajudar os alunos a uma melhor consciencialização da progressão da sua aprendizagem. Para outras, ainda, a decisão voluntária de participar neste Projeto será uma conjugação destas e/ou de outras razões.

A decisão de adesão ao Projeto é materializada na escolha do elenco de testes a realizar. Se é verdade que existem escolas que optam por uma adesão total, isto é, escolhem realizar todos os testes oferecidos (17,6% das TI EBS apresentam este comportamento, conforme referido anteriormente), as escolas restantes compõem o seu elenco de TI, estabelecendo prioridades e tirando partido de poderem *dizer*, no seu contexto específico, o que querem e/ou podem fazer⁴⁹.

Até esta fase, o estudo centrou-se na descrição e caracterização da intensidade e variedade da adesão, perspetivadas a partir de um conjunto de variáveis administrativas (tipologia de participação, natureza jurídica e localização geográfica).

Pretendeu-se nesta etapa que a análise se focasse na deteção de padrões de comportamento de escolas participantes, perante a necessidade de comporem o seu elenco de TI. Com este objetivo, recorreu-se à análise multivariada em componentes principais. Do exercício realizado, resultaram quatro componentes que aglomeram disciplinas com TI⁵⁰ cuja seleção

⁴⁹ Não serão despidiendas as razões de não escolha resultantes de dificuldades ou impossibilidades logísticas, considerando que a implementação do Projeto, no cumprimento dos procedimentos e regras estipulados, acarreta um considerável investimento, implicando afetação de meios humanos e materiais (a esta questão faz referência o *Relatório Projeto Testes Intermédios 2012* (GAVE, 2013b), quando, na sua introdução, se afirma “Embora sem significado estatístico, uma possível explicação para esta variação [decrécimo de inscrições – entre 0,3% e 5,2% - em testes de disciplinas não sujeitas a avaliação externa, entre outras três razões enunciadas, [reside em] dificuldades logísticas decorrentes da implementação do elenco completo de TI” (p.7).

⁵⁰ Atendendo ao objetivo desta análise, foram consideradas 18 disciplinas que correspondem discriminadamente a cada um dos testes que cada disciplina ofereceu para um ano de escolaridade. Desta forma,

é suscetível de corresponder a padrões de escolha (Quadro 14), ou seja, agregam subconjuntos de variáveis (de disciplinas TI) mais correlacionadas entre si e menos associadas a variáveis de outros subconjuntos.

A solução encontrada – quatro componentes retidas – explica mais de 70% da variância total, estando assegurados os requisitos necessários à realização da ACP⁵¹.

Na primeira componente, predominam os TI que correspondem a *Opções abrangentes*, isto é, estão agregados TI de disciplinas que não apresentam correspondência em provas de avaliação externa – História, Geografia, Ciências da Natureza, Inglês, Física e Química (todas de 9.º ano) com *loadings* superiores a 0,8 – ou que não apresentam exames nacionais obrigatórios – Filosofia de 10.º e de 11.º ano. Esta última disciplina, com os testes de 10.º e de 11.º ano, é a que tem menos força no conjunto elencado (*loadings* abaixo de 0,7). O conjunto de disciplinas compõe um padrão de escolha em que o denominador comum se define pela seleção de TI ser guindada por outras razões que não exclusivamente a de utilização destas provas como treino para a realização de provas de avaliação externa. Esta componente é a que mais explica a variância total dos dados originais (mais de 1/3 da variância total – 35,8%).

A segunda componente agrega um subconjunto de TI que, em comum, apresentam o facto de serem todos de ensino secundário, terem testes em mais do que um ano de escolaridade (testes de 10.º e 11.º anos), de terem correspondência em provas de exame nacional e, ainda, de serem da área das ciências. Registe-se que este fator não comporta Português (12.º ano) ou Filosofia (10.º e 11.º anos). Por outro lado, é de assinalar que os testes que definem mais vincadamente esta componente (os quatro primeiros, com *loadings* entre 0,861 e 0,831) correspondem a testes de anos de escolaridade não coincidentes com o ano de exame

pretendeu-se isolar a escolha que recaiu numa disciplina em função do ano de escolaridade para o qual o teste foi disponibilizado – por exemplo, Biologia e Geologia_10.º ano e Biologia e Geologia_11.º ano são, para este efeito, consideradas duas disciplinas.

⁵¹ Foram assegurados os requisitos necessários para a realização de uma Análise em Componentes Principais, a saber: a presença de variáveis métricas; uma adequada dimensão da amostra (550 casos); existência de multicolinearidade (correlações) entre as variáveis de input (KMO = 0,849) e, finalmente, a matriz de correlações não é uma matriz de identidade ($X^2_{(153)} = 7349,152$; $P < 0,001$).

nacional⁵². Parece haver nesta agregação, denominada *Programa completo*, um padrão de escolha em que a aposta resultará da vontade de tão cedo quanto possível no ciclo de estudos de cada disciplina expor os alunos à experiência de realização de TI. Esta componente integra também um *loading* para a Matemática A do 12.º ano (*loading* de 0,560), no entanto, esta disciplina é mais determinante na quarta componente.

A terceira componente, *Provas no 3.º CEB*, reúne os TI que, pertencentes ao 3.º ciclo, são de disciplinas com provas de avaliação externa, com especial poder agregador do TI de Matemática de 9.º ano. Isola-se, assim, um outro padrão: a escolha cujo denominador principal é a realização de TI em disciplinas de 9.º e de 8.º ano com provas finais nacionais.

Por fim, a quarta componente, *Exames no ES*, refere-se a disciplinas do ensino secundário com TI em anos de escolaridade coincidentes com os exames do ensino secundário. Neste conjunto, destacam-se as disciplinas de Português e de Matemática A (ambas do 12.º ano), acompanhadas por outras duas disciplinas cujos TI coincidem com o ano de realização de exame (Biologia e Geologia e Física e Química A, ambas do 11.º ano).

A percentagem de variância explicada pelas componentes 1 e 2 (35,80% e 20,90%, respetivamente, num total de 72,8%) faz destas componentes relevantes para a configuração de padrões de adesão.

⁵² 10.º ano para Física e Química A e Biologia e Geologia, disciplinas com exame no final do 11.º ano; e 10.º e 11.º anos para Matemática A, disciplina trienal, com exame no final do 12.º ano.

Quadro 14 – Projeto TI – Padrões de seleção
Análise de Componentes Principais

TI, disciplinas e anos de escolaridade (2010-2013)	Componentes			
	Opções abrangentes	Programa completo	Provas no 3.º CEB	Exames no ES
Hist_9	0,882			
Geo_9	0,877			
CN_9	0,849			
Ing_9	0,820			
FQ_9	0,816			
Filos_11	0,646			
Filos_10	0,613			
Mat A_10		0,861		
FQ A_10		0,859		
Mat A_11		0,846		
BG_10		0,831		
BG_11		0,721		0,517
FQ A_11		0,706		0,481
Mat_9			0,833	
Port_9			0,772	
Mat_8			0,631	
Port_12				0,779
Mat A_12		0,560		0,602
Percentagem de variância explicada	35,80%	20,90%	8,80%	7,30%

Nota: Valores próprios retidos das variáveis, superiores a 0,400; variância total explicada = 72,8%

A interpretação que se apresentou a propósito da extração destas componentes não pode deixar de ser lida num quadro de prudência que o rigor aconselha. Sem querer adiantar conclusões, a análise desenvolvida nas secções anteriores revelou uma adesão maciça das escolas ao Projeto, patente na quantidade de escolas participantes, de testes seleccionados, bem como na variedade de disciplinas cobertas. Neste quadro, exercícios de padronização, a partir dos dados disponíveis, são naturalmente pouco discriminatórios. As combinações possíveis (número de testes/seleção de disciplinas) são múltiplas, mas não serão, por natureza, substancialmente diferentes.

Se é verdade que, por exemplo, a componente que denominámos *Opções abrangentes* revela uma clara vontade de assegurar a realização daquele conjunto de TI, com aquelas características (o que decorrerá de uma vontade expressa de não perder a oportunidade de aplicar TI que poderão funcionar como instrumentos de aferição do desempenho relativamente a padrões nacionais, em áreas não sujeitas a avaliação externa), não será de

excluir que esta vontade surja no seio de escolas que apresentam taxas de adesão altas, ou seja, que optaram por fazer uma cobertura extensiva do elenco oferecido.

A análise apresentada neste trabalho não explorou outras dimensões (por exemplo, a relação entre as taxas de adesão e as componentes extraídas), não querendo deixar de aqui aflorar a problemática, a bem do rigor da investigação.

CAPÍTULO 6 – CONCLUSÃO E REFLEXÕES FINAIS

NOTA PRÉVIA

O percurso de investigação que agora se conclui assenta numa abordagem de carácter primário à avaliação de um projeto desenvolvido sob a responsabilidade de um organismo tutelado pela administração central, na área da educação, e, muito concretamente, no âmbito da avaliação das aprendizagens.

O presente estudo, intitulado *Para uma avaliação do Projeto Testes Intermédios*, pretende ser exatamente isso, um contributo para a avaliação de um projeto inscrito numa política pública de educação, partindo da necessidade de recolher, sistematizar, descrever e caracterizar um conjunto de informação de base, para que outras análises possam ser desenvolvidas.

O objetivo de descrição, análise e interpretação da adesão das escolas ao Projeto, que sustenta o estudo, orientou a abordagem e o percurso, constituindo a sua concretização a base para traçar caminhos futuros. São os resultados apurados a plataforma das conclusões que agora se apresentam e são estes mesmos dados que permitem, de forma consolidada, ajudar a desenhar outros rumos de investigação.

Em suma, de acordo com o que se propôs inicialmente, o estudo configura um contributo *para a* avaliação do Projeto e não *a* avaliação do Projeto. Sendo parcial e limitada, não deixa de ser uma contribuição estrutural, sobre a qual se poderão vir a desenvolver outras, que produzirão ou ajudarão a produzir julgamentos de valor e de mérito sobre a intervenção proposta pelo Ministério da Educação e Ciência, através do IAVE, I.P..

A relevância e a pertinência do trabalho, justificadas oportunamente na secção que lhes está dedicada, assentam naquilo que se poderia designar pelo carácter estrutural do presente estudo – a necessidade de prestação de contas e a não menos relevante necessidade de produzir informação e conhecimento capazes de melhorar a intervenção política. A este

carácter estrutural, junta-se a atualidade que decorre da recente decisão de descontinuar o Projeto TI, anunciada em setembro de 2014. Este facto justifica a realização de um balanço, para que na desejada consolidação do *framework* nacional de avaliação⁵³ comecem a ficar mais claros, porque mais informados, o racional e a narrativa sobre como a avaliação se relaciona com o desenvolvimento do sistema educativo português, sabendo como cada elemento desse *framework* poderá contribuir para a melhoria das práticas escolares.

Quase dez anos de desenvolvimento do Projeto Testes Intermédios, as suas características, a sua dimensão e abrangência consubstanciam-no como um dos elementos constituintes do referido *framework* avaliativo e, portanto, como responsável por alterações ocorridas no âmbito do que se identifica como cultura de avaliação, no quadro de muitas outras alterações ocorridas ao longo dos últimos catorze anos.

É nesta perspetiva que, na existência e na descontinuidade do Projeto, a sua avaliação é abordada, sendo que o triénio em análise (2010-2013), pelas características que apresenta (diversificação da oferta disciplinar, em grande parte decorrente de cobertura de disciplinas do 3.º CEB, relativa estabilidade interanual de testes em quantidade e variedade, além de outras questões de carácter administrativo), se constituiu como uma fonte apta para gerar informação quantitativa e qualitativa relevante.

SÍNTESE DOS PRINCIPAIS RESULTADOS

A matriz de investigação desenhada definiu como objeto de estudo a *Adesão das escolas ao Projeto TI*, abrangendo quatro dimensões:

- Elenco de TI a aplicar em cada ano.
- Evolução da adesão das escolas.
- Distribuição da adesão.
- Padrões de adesão.

⁵³ Recupera-se aqui a remissão feita anteriormente para o relatório da responsabilidade de Santiago *et al.* (2012).

A análise de dados realizada permite, nesta fase, reunir um conjunto de conclusões relevantes para a resposta às questões de investigação formuladas⁵⁴. Assim, e tendo sempre por referência o triénio analisado, podemos concluir:

1. Regista-se uma participação expressiva de escolas no Projeto: um total de 1378 unidades participantes, estando uma clara maioria (88,2%, correspondentes a 1215 escolas) presente nas três edições (2010/2011, 2011/2012 e 2012/2013).
2. A adesão média (calculada com base no número de testes selecionados anualmente para composição do elenco por referência ao número de testes oferecidos) mantém-se superior a 50% em todos os anos, com valores que revelam fidelização ao Projeto.
3. A participação das escolas fez-se independentemente da sua natureza jurídica: estiveram representadas escolas públicas e privadas, numa proporção idêntica à distribuição da rede escolar nacional para o período em análise.
4. No que respeita à adesão média, as escolas privadas apresentam valores mais elevados do que as públicas – em qualquer um dos anos, superiores a 60%.
5. O Projeto adquiriu expressão nacional consistente, com todas as regiões a serem representadas através de escolas participantes, havendo ainda representação continuada das EPE.

⁵⁴ Reproduzem-se as questões de investigação formuladas:

- Como se poderá caracterizar o universo de escolas que aderiram ao Projeto TI no triénio 2010-2013?
- Como se poderá caracterizar a natureza da adesão das escolas ao Projeto?
 - Qual a quantidade e variedade de TI escolhidos pelas escolas?
 - Que padrões de participação (traduzidos na constituição do elenco de TI) poderão ser identificados?

6. Das 31 unidades geográficas consideradas⁵⁵, a maioria apresenta médias de adesão superiores à média global, chegando algumas a atingir valores entre 60% e 80% de adesão nos três anos.
7. As taxas de cobertura do Projeto TI (percentagem de escolas que aderiram ao Projeto) mostram valores muito próximos de 100% (oscilando, nos três anos, entre 94,8% e 99,4%), o que é revelador da adesão maciça das escolas, acompanhada de uma constância na vontade de participação.
8. Os TI mais populares (com a popularidade a ser medida pelas taxas de adesão anuais) são os das disciplinas que apresentam correspondência em provas de avaliação externa (de carácter obrigatório), tanto no 3.º CEB como no ES. Nestes, as taxas ascendem, por regra, a valores superiores a 90%, o que revela uma adesão expressiva e, uma vez mais, consistente, perante a oferta.
9. Os TI de disciplinas que não têm provas de avaliação externa ou cujas provas apresentam carácter opcional têm taxas de adesão mais baixas, mas sempre superiores a 45%, na maioria dos casos a oscilarem entre 50% e 65% (atingindo mesmo 71,1%, no caso específico da disciplina de Filosofia, teste de 11.º ano, em 2011/2012). Neste grupo de TI, e para as disciplinas do 3.º CEB, destaca-se a disciplina de Físico-Química (9.º ano), com o TI mais popular em todos os anos do período analisado.
10. Quando a análise se centra na denominada tipologia de participação Escolas TI EBS, os dados relativos à adesão média, no triénio, ascendem a 80,7%, mantendo-se a sobrerrepresentação da adesão média das escolas privadas relativamente às públicas.
11. No que respeita à distribuição geográfica da adesão, continuando a considerar as TI EBS, regista-se novamente uma extensiva cobertura nacional, acompanhada de taxas de adesão significativamente elevadas. Neste universo de escolas, a região Centro

⁵⁵ Em rigor, 30 unidades correspondentes a NUTS III e as EPE.

(NUTS II) reúne a esmagadora maioria das sub-regiões que apresentam as taxas mais elevadas de adesão, com valores acima de 85% (correspondentes a um número médio de inscrições por escola superior a 45,1 – para uma oferta, no triénio, de 53 testes). Mesmo no extremo oposto a este *ranking*, os valores médios de adesão nunca descem de 70%.

- 12.** Consideradas estas 550 escolas (as TI EBS), regista-se um comportamento de seleção de quantidade e de variedade de TI assinalável: 65,8% das escolas optaram por realizar, no triénio, 75% ou mais do número de TI oferecidos, sendo que destas aproximadamente 18% realizaram 100% da oferta, num total de 53 TI nos três anos. Por outro lado, não se registam situações de escolas que tenham optado por realizar menos de 25% do número de testes propostos. No que respeita à variedade, são igualmente significativos os dados apurados: 71,6% das escolas optaram por realizar pelo menos uma prova de todas as disciplinas disponibilizadas, sendo que 319 escolas quiseram ver coberto o elenco disciplinar disponível.
- 13.** A existência de um padrão de adesão que apresenta a percentagem mais elevada de variância explicada (a denominada componente *Opções abrangentes*) significará que, na composição do elenco, estas escolas guindam as decisões relativas à composição do seu elenco de TI por razões que não se prendem exclusivamente com a utilização dos testes como antecâmaras das provas de avaliação externa. Ou melhor, guindam as decisões relativas à composição do elenco TI tendo como objetivo cobrirem as disciplinas que têm TI, mesmo não havendo provas de avaliação externa.

REFLEXÕES A PARTIR DA SÍNTESE DE RESULTADOS

A síntese de resultados que se apresentou suporta indiscutivelmente a significativa dimensão e consistência de adesão das escolas ao Projeto TI. Não é de mais reiterar a opcionalidade que sempre presidiu à participação e à constituição do elenco anual de testes. A decisão de participar, e de como participar, constituía uma prerrogativa que cada uma das escolas, em

cada ano, usava, de acordo com um conjunto de critérios que definiria, no contexto particular do seu projeto educativo e no seio da sua comunidade educativa, após conhecimento e análise da proposta anualmente apresentada pelo IAVE, I.P..

Àquele direito de escolha, juntava-se um outro de natureza igualmente relevante: cumpridos os procedimentos de aplicação dos testes e de devolução das grelhas de classificação, as implicações atribuídas aos seus resultados, ou melhor, a função (ou funções) que a informação avaliativa gerada pela realização destes instrumentos de avaliação desempenhava(m) na avaliação interna dos alunos era matéria da exclusiva responsabilidade da escola.

Essas duas dimensões, entre outros aspetos que mencionaremos, constituíam um traço identitário do Projeto que dá significado às taxas de participação: perante um exercício de liberdade de escolha, as escolas aderiram ao que lhes foi proposto, fizeram-no consistentemente, de forma comprometida, geriram internamente as suas opções. Fizeram-no independentemente da sua natureza pública ou privada, e a sua participação teve expressão nacional. De facto, a opcionalidade conviveu com a popularidade.

Recuperamos as finalidades enunciadas pelo Projeto, para podermos prosseguir o exercício de interpretação dos resultados obtidos.

“[a)] permitir a cada professor aferir o desempenho dos seus alunos por referência a padrões de âmbito nacional, [b)] ajudar os alunos a uma melhor consciencialização da progressão da sua aprendizagem e, [c)] complementarmente, contribuir, sempre que aplicável, para a sua progressiva familiarização com instrumentos de avaliação externa” (GAVE, 2012, p. 2).

Estas finalidades eram clarificadas, no documento citado, através do reforço do carácter eminentemente formativo dos TI, a par da reafirmada necessidade de aferição dos desempenhos e de regulação da aprendizagem, a partir de uma reflexão sustentada pela análise do processo de resposta dos alunos e pelos resultados atingidos.

Ora, este quadro referencial era constituído por um conjunto vasto e ambicioso de objetivos que terão presidido, teoricamente, à decisão das escolas. O peso que cada um dos objetivos

assumiu na construção da decisão é matéria sobre a qual os dados tratados não geram diretamente informação.

Porém, numa primeira perspetiva de abordagem, a prioridade dada às disciplinas com TI que tinham correspondência em provas de avaliação externa, conforme constatado anteriormente, permitirá identificar a finalidade de *progressiva familiarização com instrumentos de avaliação externa* como a que assumiu um relevo especial no processo de decisão.

Uma análise desta primeira inferência não poderá deixar de convocar outras dimensões do que será esta *familiarização*. A vontade de exposição dos alunos a provas de avaliação produzidas centralmente pela entidade que tem a responsabilidade de também conceber as provas finais e exames e o facto de a aplicação ser síncrona, nacional, e com procedimentos que podem, em última análise, replicar os do *dia do exame* são fatores que terão peso na escolha prioritária destas áreas disciplinares. No entanto, não se crê que esta dinâmica esgote as possibilidades de explicação da adesão.

De facto, as elevadíssimas taxas de cobertura dos TI destas disciplinas (com valores, em regra, superiores a 90%) obrigam a olhar para este processo como uma declaração de vontade (de desejo) quase universal de, através dos TI, conseguir obter melhores resultados na avaliação externa. Nesta perspetiva, os TI, enquanto *antecâmaras* das provas finais e dos exames, terão congregado persistentemente uma vontade de melhoria, mesmo para todas aquelas escolas que não viram, ao longo dos anos, melhorar os seus desempenhos em avaliação externa.

E essa persistência conduz-nos, ainda, a outras dimensões de análise. Muito provavelmente, as escolas terão reconhecido a estes instrumentos de avaliação outras potencialidades: terão reconhecido nos TI formas privilegiadas de aferir o desempenho dos seus alunos face a padrões nacionais. Nesta perspetiva, terão entendido que os TI constituíam produtos *tecnologicamente robustos*⁵⁶, ou mais robustos do que outros *testes de papel e lápis*, e,

⁵⁶ Recupera-se aqui a conceção expressa por Madaus, Russell & Higgins (2009) de que a construção de um teste é *tecnologia*.

portanto, mais aptos para avaliar com qualidade (de forma mais válida e mais fiável) as aprendizagens dos alunos. Ou tão-somente terão considerado que a representação de currículo que os TI apresentavam era a oficial, a legítima, porque chancelada pelo MEC e/ou porque desenvolvida por uma instituição com larga experiência em construção de testes.

E a essa legitimidade acrescida no estabelecimento do padrão de desempenho, juntava-se a possibilidade de, com os seus resultados, a escola se posicionar perante os resultados regionais e nacionais. As escolas terão, de facto, procurado também isto: um *standard* que permitisse aferir o seu trabalho interno e posicioná-lo no contexto externo.

A procura de uma posição relativa no contexto nacional e, fundamentalmente, a aferição do desempenho interno, através de um conjunto de provas de conceção externa⁵⁷, não são, naturalmente, um exercício que se esgota em si. Só faz sentido se a análise gerar *feedback* utilizável na escola. E se esta é uma área de complexa abordagem, e sobre a qual, no âmbito deste trabalho, não é possível desenvolver análise, será fácil conceder, assente na confiança na vontade de melhoria, que alguma coisa, muitas coisas, coisas diversas, as escolas terão feito em função da sua participação no Projeto TI.

A assunção feita anteriormente suporta a convicção de que as finalidades de (i) aferição de desempenho face a padrões nacionais e de (ii) regulação das aprendizagens (recorrendo à função eminentemente formativa que o Projeto preconizava) também terão, quando não alavancado a participação, sido alcançadas pela participação em si mesma.

Nesta fase, e a bem do enquadramento inicialmente feito, designadamente a propósito da avaliação inscrita no relatório sobre Portugal (Santiago *et al.*, 2012), recorda-se: “The majority of schools use them [os TI] as summative assessment instruments.” (p. 54). Esta conclusão, que a equipa responsável pelo Relatório produz em 2012, com base em informação recolhida no ano anterior⁵⁸, identifica aquilo que pode ser entendido como um aproveitamento menos valioso da informação gerada pelos TI. Efetivamente, é à avaliação

⁵⁷ A desagregação de informação que as grelhas de classificação dos TI apresentavam permitia a análise de desempenho a vários níveis (item, grupo de itens, distribuição geral de resultados, entre outros indicadores).

⁵⁸ A visita da equipa de avaliação ocorreu entre fevereiro e março de 2011.

de carácter formativo, como já foi referido, que se reconhece a vocação para o apoio ao processo de ensino e de aprendizagem. Esta opção por uma utilização que privilegia o balanço, a síntese, do que o aluno sabe, que avalia aquilo que aprendeu, sem estar, à partida, vocacionada para estar ao serviço das aprendizagens, para as aprofundar e ajustar, será, assim, uma apropriação menos rica das potencialidades dos TI. Mas esta constatação, que se toma por correta e a corresponder ao comportamento maioritário das escolas, não faz dos TI instrumentos com implicações negativas ou não aptos para apoiar o processo de elevação de *standards*. Só consubstancia usos, eventualmente, menos potenciadores da melhoria.

Aquela constatação não exclui a possibilidade de as escolas se terem apropriado da informação gerada pelos TI de forma não unívoca, isto é, de os TI terem produzido informação de natureza sumativa e formativa. Efetivamente, como vimos anteriormente, a utilização de informação avaliativa para suportar a atribuição de uma classificação, que pretende informar sobre o nível de desempenho atingido, não exclui outras utilizações da informação recolhida, mais dirigidas ao apoio e à reorientação do processo do ensino e de aprendizagem e, portanto, de vocação formativa.

Aliás, a eventual menor capacidade de o Projeto induzir mudanças profundas e estruturais das práticas avaliativas não será, na essência, muito diferente daquilo que faz parte de uma *agenda por cumprir* em termos avaliativos. Fernandes (2014), referindo-se à atualidade, afirma:

avaliar para classificar, para selecionar ou para certificar continuam a ser preocupações marcantes nas práticas de avaliação. Por isso, as políticas educativas deverão dar prioridade à melhoria das aprendizagens nas salas de aula, isto é, ao desenvolvimento de práticas sistemáticas de avaliação formativa (p. 267).

Numa segunda perspetiva de abordagem a estas reflexões finais a partir dos resultados apurados, centra-se a análise na escolha dos outros TI, dos que marcam o elenco do triénio 2010-2013, por serem de disciplinas sem provas de avaliação externa, disciplinas inéditas no Projeto e cuja oferta se descontinuou logo em 2013/2014.

Se é verdade que se registou, no período em análise, uma preponderância da escolha de TI de disciplinas com provas de avaliação externa, não deixa de ser igualmente relevante recuperar, neste momento, o padrão de adesão que denominámos *Opções abrangentes*. Além das considerações que aqueles dados já nos mereceram, será legítimo inferir que estas escolhas terão resultado da vontade de assegurar a realização de provas de avaliação concebidas centralmente sobre áreas curriculares não sujeitas a avaliação externa. Equivale isto a dizer que a escolha decorrerá da necessidade que as escolas terão sentido de regulação do currículo, de aferição de conhecimentos dos alunos relativamente a padrões externos em áreas sobre as quais não há regulação, e, ainda, da necessidade de moderação da avaliação interna, ou mesmo de um processo de validação da avaliação interna através da avaliação externa⁵⁹. Complementarmente, esta adesão poderá ainda oferecer leituras relacionadas com a utilização destes TI como forma de motivar os alunos para o investimento em áreas que poderão estar a ser secundarizadas, por efeito do estreitamento de currículo resultante da concentração de avaliação externa, no caso do 3.º CEB, em Português e Matemática.

Este último aspeto ganha especial relevância na medida em que a oferta de instrumentos como os TI nestas outras áreas disciplinares (não sujeitas a provas de avaliação externa) terá contribuído para minorar os impactos não desejados da aplicação de programas de avaliação em larga escala. A existência de TI, integrados num programa nacional continuado, era a forma de ver mais controlados os efeitos colaterais negativos que abordámos em sede de revisão de literatura. Nessa perspetiva, a possibilidade de mitigar esses efeitos foi também descontinuada. Os serviços centrais deixaram de recolher informação sobre a concretização curricular de várias áreas disciplinares e as escolas terão deixado de poder aferir os seus desempenhos e de, por esta via, prestar contas.

É legítimo, nesta fase, e mesmo antes de prosseguir para uma última perspetiva de análise, registar que as potencialidades que o Projeto ofereceu e que as escolas terão desenvolvido não terão relevado o suficiente para contrariar a intenção de descontinuar o Projeto nos moldes até então vigentes, registando-se a dificuldade em compreender a decisão pela inexistência de evidências públicas de base empírica ou outras que a sustentem.

⁵⁹ No caso concreto, através de instrumentos de avaliação que têm um carácter híbrido, mas que, ainda assim, apresentam esta função.

Abordadas as duas perspetivas anteriores, mais parcelares, avança-se para uma terceira e última, mais global. Explora-se a interpretação dos resultados apurados no estudo à luz do que inicialmente se designou por uma *já longa etapa de transformação global* (iniciada em 2000), que integra, entre outras, intervenções de fundo no domínio da avaliação das aprendizagens.

Nesse caminho que o sistema educativo português tem vindo a percorrer, há uma unanimemente considerada evolução naquilo que habitualmente se designa por cultura de avaliação, para a qual terão contribuído, entre muitíssimos outros aspetos, o enquadramento legal do sistema de avaliação do ensino básico e secundário⁶⁰, a criação do Instituto de Inovação Educacional, a participação em estudos internacionais, a criação do GAVE, a existência de provas globais, a introdução de provas de avaliação externa no ensino básico (*low-stakes* e *high-stakes*) e a sua continuidade no ensino secundário (*high-stakes*), a publicitação dos resultados da avaliação externa, bem como o sistema de avaliação externa de escolas (no que diretamente se relaciona com a avaliação de alunos).

Os fatores elencados, a título meramente exemplificativo, terão contribuído para a modificação gradual da atitude da comunidade educativa, dos investigadores, da comunicação social, dos responsáveis políticos e da sociedade, em geral, relativamente aos resultados escolares.

Para esta alteração, que se defende pertencer ao âmbito de uma nova *cultura de avaliação*, porque envolve modificações em formas de organização do trabalho e em códigos de relacionamento interpessoal dentro de uma dada comunidade, terá contribuído igualmente a existência durante uma década (2005-2014) do Projeto TI.

Nesta perspetiva, o Projeto terá tido efeitos, por um lado, na forma como as escolas se (re)organizaram para responder aos desafios que este lhes oferecia, por outro, na maneira como os resultados, na sua função de prestação de contas, foram apropriados pelos pais e

⁶⁰ Com especial destaque para o carácter fundador do Despacho Normativo n.º 98-A/92, de 19 de junho, que lança os fundamentos de um conjunto de princípios e de conceções que se mantêm até à atualidade.

encarregados de educação e por outros *stakeholders*, e, ainda, numa terceira vertente, na informação que se produziu do sistema para o sistema, para a instituição que geriu o Projeto e que tratou dados e os devolveu às escolas.

Centra-se a análise na primeira vertente identificada – a dos efeitos escolares, junto dos professores e do seu trabalho na escola, porque será esta a perspetiva que mais poderá usufruir de leituras a partir dos dados que se trataram.

De facto, o processo que levou à adesão das escolas e à constituição do elenco anual, como já se viu, envolveu 1215 escolas no triénio e, por maioria de razão, envolveu um número elevadíssimo de alunos e de professores. Os professores, através das suas estruturas de representação, terão contribuído para as decisões e os alunos terão inevitavelmente visto a sua rotina escolar marcada pela realização das provas.

Um conhecimento, mesmo que genérico, da organização escolar permite afirmar que as escolas viram as suas rotinas de trabalho curricular e de avaliação (além de outros aspetos de natureza diferente) alteradas por esta intervenção.

A planificação de um ano escolar que incluía a realização de TI obrigou à articulação do trabalho entre professores, à existência de trabalho em equipa, à partilha e à reflexão. Terá obrigado a revisitar os documentos de referência curricular e a questionar opções. Nuns casos, terá sido mais uma oportunidade de efetivo trabalho científico e didático em equipa, noutros, terá constituído a oportunidade para não mais adiar as potencialidades desse mesmo trabalho em equipa. Em qualquer uma das situações, os professores terão experimentado outras formas de construir testes, outras formas de desenhar critérios de classificação, ter-se-ão apropriado da forma de os aplicar, terão discutido, trabalhado em conjunto, e em alguns casos classificado conjuntamente⁶¹. A experiência, marcada pelos compromissos assumidos aquando da decisão de participação e pelo escrutínio a que processos nacionais desta dimensão sempre constituem, terá igualmente conduzido ao desenvolvimento de práticas de leitura dos resultados, através de exercícios de interpretação dos dados estatísticos gerados

⁶¹ As práticas de trabalho colaborativo entre professores são reconhecidas como um dos fatores determinantes na melhoria do desempenho dos sistemas educativos. Esta também é, naturalmente, matéria abordada nos relatórios OCDE já anteriormente mencionados (Santiago *et al.*, 2012, 2013).

por cada um dos testes e pelos relatórios nacionais publicados. As aprendizagens terão tido, como é natural, repercussões mais amplas, que ultrapassarão o âmbito intrínseco do Projeto.

Por outro lado, para os alunos (para muitos e em diferentes anos do seu percurso escolar, se considerarmos a diversidade disciplinar oferecida no triénio 2010-2013), a participação no Projeto terá dado a experimentar outros modelos de instrumentos de avaliação, a possibilidade de perspetivarem o seu percurso de aprendizagem a partir de avaliações de carácter externo, de âmbito nacional, com um olhar que lhes permitiu sair da especificidade do que acontece na sala de aula de cada um e de lá voltar com um olhar criticamente informado. Terão experimentado a urgência de consultar critérios de classificação publicados pelo IAVE no próprio dia da realização do teste e de, com a proficiência possível, desenvolverem processos de autoclassificação. Terão tido condições para serem atores, mais ou menos envolvidos, do processo acima descrito, numa representação que congregava, além da prestação individual, a prestação da turma e da escola.

Toda a dinâmica descrita apontará para o despontar ou o desenvolvimento, na escola, de um conjunto de padrões de comportamento de valor acrescido diretamente relacionados com a existência do Projeto. A instância de moderação, até agora induzida através destes instrumentos de avaliação, e que afinal poderá contribuir para explicar a adesão generalizada e consistente ao Projeto TI, terá sido uma instância de moderação do ensino, da aprendizagem e da avaliação.

Esta instância de moderação terá sido porventura o contributo mais estrutural que o Projeto levou às escolas. De que forma este contributo terá condições para frutificar é matéria que necessitará de acompanhamento e avaliação.

NOTA FINAL

A forma de fazer evoluir, de modo sustentado porque informado, os programas de política pública e, no caso concreto, o juízo de mérito e de valor sobre um projeto desta natureza é uma tarefa árdua e de relevante complexidade.

A avaliação da qualidade é um exercício de representação, é multifacetada, nunca completamente representada, é a que é provável, é suscetível ao olhar do observador e o processo de avaliação deve ser considerado um *caso em aberto* (Stake & Schwandt, 2006).

Ter isto em mente é ter a noção da importância de como se avalia, ter a responsabilidade de cumprir processos rigorosos, na vontade de que a investigação seja uma contribuição para melhor conhecer, melhor compreender e melhor suportar a decisão.

O estudo que agora se conclui pretendeu ser um contributo para melhor conhecer e melhor compreender, podendo vir a sustentar outras perspetivas de investigação.

GLOSSÁRIO

TAXA DE ADESÃO

Percentagem média de TI realizados, face ao elenco disponibilizado no triénio 2010-2013.

TAXA DE COBERTURA

Percentagem de escolas participantes no Projeto TI relativamente ao universo de escolas do país.

BIBLIOGRAFIA

- ARG (Assessment Reform Group) (1999) *Assessment for learning. Beyond the black box*. University of Cambridge: Assessment Reform Group.
- ARG (Assessment Reform Group) (2002). *Assessment for learning. 10 principles*. University of Cambridge: Assessment Reform Group. Consultado em 2 de janeiro de 2015, em <<http://www.aaria.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf>>
- ARG (Assessment Reform Group) (2003). *The role of teachers in the assessment of learning*. University of Cambridge: Assessment Reform Group.
- Beaton, A., Postlethwaite, T., Ross, K., Spearritt, D., Wolf, R. (2000). *The Benefits and limitations of international educational achievement studies*. Paris: International Academy of Education.
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: GL Assessment.
- Black, P., & Wiliam, D. (2006). The reliability of assessments. In J. Gardner (Ed.), *Assessment and learning*, pp. 119-133. London: Sage.
- Black, P. (1998). *Testing: friend or foe? Theory and practice of assessment and testing*. London: Falmer.
- Butt, G. (2010). *Making Assessment Matter*. London: Continuum.
- Carvalho, H., Ávila, P., Ramos, M. (2010) *Análise de dados em Ciências Sociais. Iniciação ao SPSS*, Lisboa: ISCTE (em PDF).
- Casassus, J. (2009) Uma nota crítica sobre a avaliação estandardizada: a perda da qualidade e a segmentação social. *Sísifo: Revista de Ciências da Educação*, 2009, pp. 71-79. Consultado em 12 de novembro de 2013, em <<http://www.sisifo.fpce.ul.pt>>.
- Donaldson, S. (2001). Mediator and moderator analysis in program development. In S. Sussman (Ed.), *Handbook for program development for health behaviour research and practice*, pp. 470-496. Newbury Park: Sage.
- Esteban, M. T. (2009). Avaliação e fracasso escolar: questões para debate sobre a democratização da escola. *Revista Lusófona de Educação*, (13), 2009, pp. 123-134. Consultado em 15 de setembro de 2014, em <http://www.scielo.gpeari.mctes.pt/scielo.php?script=sci_arttext&pid=S1645-72502009000100008&lng=pt&tlng=pt>
- Fernandes, D. (Coord.) (n.d.). Instrumentos de avaliação: diversificar é preciso. In *Pensar a avaliação, melhorar a aprendizagem*. Lisboa: IIE.

- Fernandes, D. (2005). *Avaliação das aprendizagens: desafios às teorias, práticas e políticas*. Lisboa: Texto Editores.
- Fernandes, D. (2010). Acerca da articulação de perspectivas e da construção teórica em avaliação educacional. In M. T. Esteban e A. J. Afonso (Orgs.), *Olhares e interfaces: Reflexões críticas sobre a avaliação*, pp. 15-44. São Paulo: Cortez.
- Fernandes, D. (2011). Avaliação de programas e projetos educacionais: das questões teóricas às questões das práticas. In D. Fernandes (Org.), *Avaliação em educação: Olhares sobre uma prática social incontornável*, pp. 185-208. Pinhais, PR: Editora Melo.
- Fernandes, D. (2014). Avaliação das aprendizagens e políticas educativas: o difícil percurso da inclusão e da melhoria. In M.L. Rodrigues (Org.), *40 anos de políticas de educação em Portugal: a construção do sistema democrático de ensino*, pp. 231-268. Lisboa: Edições Almedina.
- GAVE (Gabinete de Avaliação Educacional) (2008). *Projeto Testes Intermédios – 2008/2009. Informação n.º 1*. Lisboa: GAVE – Ministério da Educação (em PDF).
- GAVE (Gabinete de Avaliação Educacional) (2012). *Projeto Testes Intermédios – 2012/2013. Informação n.º 1*. Lisboa: GAVE – Ministério da Educação e Ciência. Consultado em 14 de junho de 2013, em <<http://www.gave.min-edu.pt/np3/2.html>>
- GAVE (Gabinete de Avaliação Educacional) (2013a). *Projeto Testes Intermédios – 2012/2013. Informação n.º 6 Português – 9.º ano – 07.fev.2013. Resultados nacionais e regionais*. Lisboa: GAVE – Ministério da Educação e Ciência. (em PDF).
- GAVE (Gabinete de Avaliação Educacional) (2013b). *Relatório Projeto Testes Intermédios 2012*. Lisboa: GAVE – Ministério da Educação e Ciência. Consultado em 14 de janeiro de 2015, em <<http://iave.pt/np4/113.html#2>>
- Gardner, J. (Ed.) (2006). *Assessment and learning*. London: Sage.
- Gipps, C. (1994). *Beyond testing: towards a theory of educational assessment*. London: Falmer.
- Guba, E. G, & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Sage Publications, Incorporated.
- Hamilton, L. S. (2003). *Assessment as a policy tool*. Review of research in education, 27, 2007, pp. 25-68.
- Harlen, W. (2006a). On the relationship between assessment for formative and summative purposes. In J. Gardner (Ed.), *Assessment and Learning*, pp. 103-117. London: Sage.
- Harlen, W. (2006b). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning*, pp. 61-80. London: Sage.
- IAVE, I.P. (Instituto de Avaliação Educativa, I.P.) (2013a). *Projeto Testes Intermédios – 2013/2014. Informação-Projeto. Informação aos Diretores, gestores de projeto e professores*. Lisboa: IAVE – Ministério da Educação e Ciência. (em PDF).

- IAVE, I.P. (Instituto de Avaliação Educativa, I.P.) (2013b). *Projeto Testes Intermédios – 1.º Ciclo do Ensino Básico – Relatório 2013*. Lisboa: IAVE – Ministério da Educação e Ciência. Consultado em 14 de janeiro de 2015, em <<http://iave.pt/np4/106.html#1>>
- IAVE, I.P. (Instituto de Avaliação Educativa, I.P.) (2014). *Projeto Testes Intermédios – 2014/2015. Informação n.º 1*. Lisboa: IAVE – Ministério da Educação e Ciência. (em PDF).
- IAVE, I.P. (Instituto de Avaliação Educativa, I.P.) (2015). *Projeto Testes Intermédios – 1.º Ciclo do Ensino Básico – Relatório 2014*. Lisboa: IAVE – Ministério da Educação e Ciência. Consultado em 14 de janeiro de 2015, em <<http://iave.pt/np4/103.html#4>>.
- JCSEE (Joint Committee on Standards for Educational Evaluation) (2011). *The program evaluation standards. A guide for evaluators and evaluation users*. Thousand Oaks: Sage.
- JNE (Júri Nacional de Exames) (2011). Exames nacionais do ensino básico 2011 – Base de dados. Consultado em 10 de setembro de 2014, em <http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>.
- JNE (Júri Nacional de Exames) (2011). Exames nacionais do ensino secundário 2011 – Base de dados. Consultado em 10 de setembro de 2014, em <http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>.
- JNE (Júri Nacional de Exames) (2012). ENEB 2012 – Base de dados das provas finais dos 2.º e 3.º ciclos do ensino básico. Consultado em 10 de setembro de 2014, em <<http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>
- JNE (Júri Nacional de Exames) (2012). ENES 2012 – Base de dados dos exames nacionais do ensino secundário. Consultado em 10 de setembro de 2014, em <<http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>
- JNE (Júri Nacional de Exames) (2013). ENEB 2013 – Base de dados das provas finais dos 2.º e 3.º ciclos do ensino básico. Consultado em 10 de setembro de 2014, em <<http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>
- JNE (Júri Nacional de Exames) (2013). ENES 2013 – Base de dados dos exames nacionais do ensino secundário. Consultado em 10 de setembro de 2014, em <<http://www.dgidec.min-edu.pt/jurinacionalexames/index.php?s=directorio&pid=33>>
- Kellaghan, T., & Madaus, G. F. (2003). External (public) examinations. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International handbook of educational evaluation*, pp. 577-600. Dordrecht: Kluwer Academic.

- Linn, R. (2001). *The design and evaluation of educational assessment and accountability systems*. CRESST/University of Colorado at Boulder. Consultado em 10 de março de 2014, em <<http://www.cse.ucla.edu/products/reports/TR539.pdf>>.
- Madaus, G., Russell, M., & Higgins, J. (2009). *The paradoxes of high-stakes testing: how they affect students, their parents, teachers, principals, schools, and society*. Charlotte, NC: Information Age Publishing.
- Marôco, J. (2014) *Análise estatística com o SPSS Statistics*. Pêro Pinheiro: ReportNumber.
- Mehrens, W. (1984). National tests and local curriculum: match or mismatch?, *Educational measurement: issues and practice*, (3), 3, pp. 9-15.
- Mehrens, W. (2002). Consequences of assessment: what is the evidence?. In G. Tindal & T. Haladyna (Eds.), *Large-scale assessment programs for all students. validity, technical adequacy and implementation*, pp. 149-177. New York: Lawrence Erlbaum.
- Pacheco, J. (2013). Avaliações internacionais e desempenho dos alunos portugueses. In CNE (Ed.) *Avaliações internacionais e desempenho dos alunos portugueses*, pp. 135-149, Lisboa: CNE. Consultado em 10 de março de 2014. em <http://www.cnedu.pt/content/edicoes/seminarios_e_coloquios/LIVRO_Avaliacoes_internacionais.pdf>.
- Pinto, J. & Santos, L. (2006). *Modelos de avaliação das aprendizagens* Lisboa: Universidade Aberta.
- Redfield, D. (2001). *Critical Issues in large-scale assessment: a resource guide*. Council of Chief State School Officers, Washington, DC: AEL, Inc., Charleston, WV.
- Rodrigues, M. L. (2010), *A Escola Pública pode fazer a diferença*. Coimbra: Edições Almedina.
- Santiago, P.; Donaldson, G.; Looney, A., & Nusche, D. (2012). *OECD Reviews of evaluation and assessment in education: Portugal*. OECD Publishing, Paris. Consultado em 15 de dezembro de 2013 em <<http://www.oecd.org/edu/evaluationpolicy>>
- Santiago, P.; Donaldson, G.; Looney, A., & Nusche, D. (2013). *Synergies for better learning: an international perspective on evaluation and assessment*. OECD Publishing, Paris. Consultado em 15 de dezembro de 2014 em <http://www.oecd.org/edu/school/Evaluation_and_Assessment_Synthesis_Report.pdf>
- Santos, L., (Org.) (2010). *Avaliar para aprender. Relatos de experiências de sala de aula do pré-escolar ao ensino secundário*. Porto: Porto Editora e Instituto de Educação, Universidade de Lisboa.
- Scriven, M. (1991). *Evaluation Thesaurus*. Newbury Park: Sage.

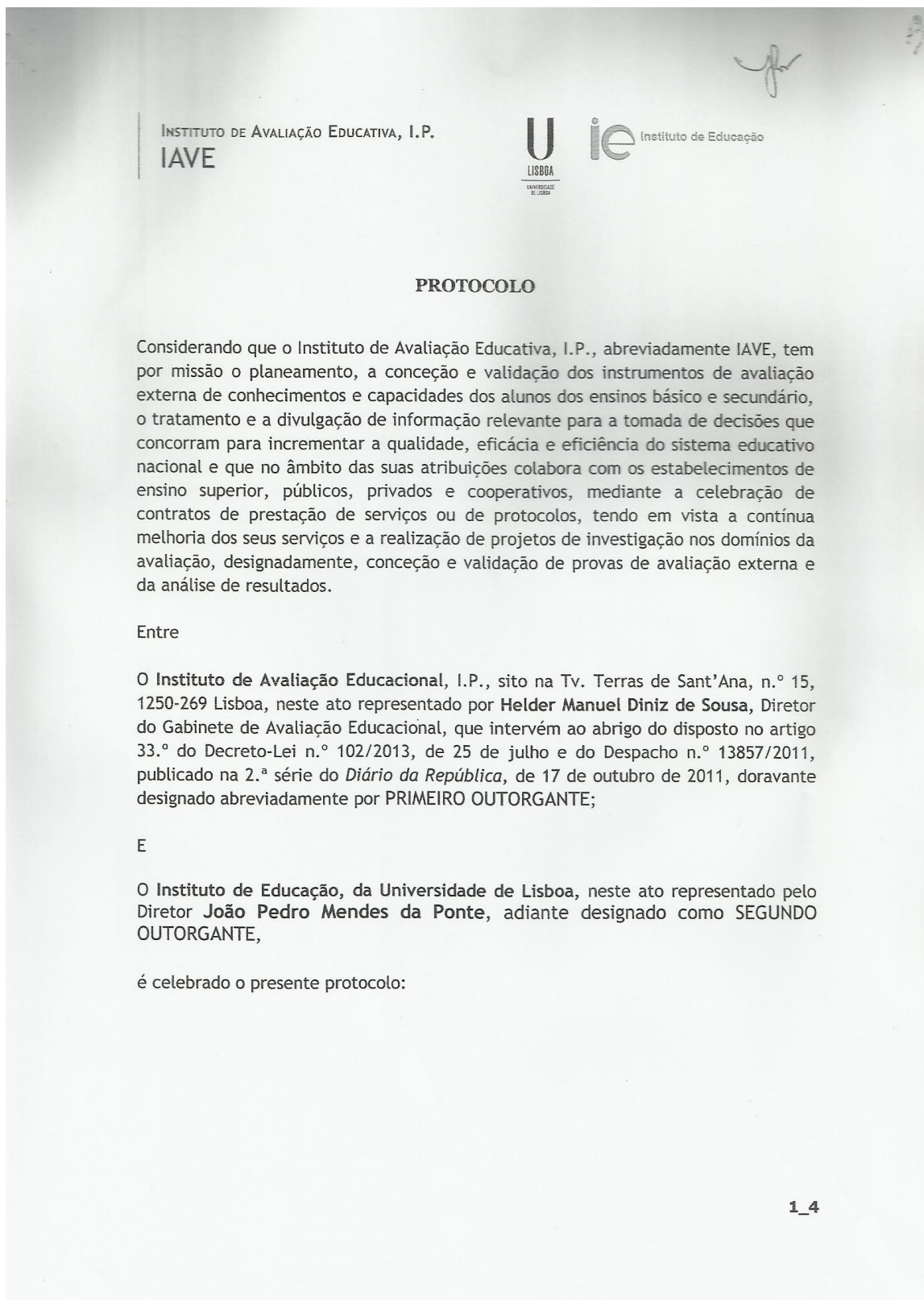
- Stake, R. & Schwandt, T. (2006). On discerning quality in evaluation. In Shaw, Greene, Mark (Eds.), *The Sage Handbook of Evaluation*. London: Sage.
- Stake, R. (2006). *Standards-based & responsive evaluation*. London: Sage.
- Stobart, G. (2008). *Testing times: the uses and abuses of assessment*. Abingdon: Routledge.
- Stufflebeam, D. & Shinkfield, A. (2007). *Evaluation theory, models, & applications*. San Francisco, CA: Jossey-Bass.

DOCUMENTOS LEGISLATIVOS:

- Despacho Normativo n.º 98-A/92, de 20 de junho. *Diário da República n.º 140/1992 – 1.º suplemento, Série I-B*. Ministério da Educação.
- Decreto-Lei n.º 299/2007, de 22 de agosto. *Diário da República n.º 161/2007 – Série I*. Ministério da Educação.
- Decreto-Lei n.º 50/2011, de 8 de abril. *Diário da República n.º 70/2011 – Série I*. Ministério da Educação.
- Portaria n.º 244/2011, de 21 de junho de 2011. *Diário da República n.º 118/2011 – Série I*. Ministério da Educação.
- Decreto-Lei n.º 139/2012, de 5 de julho. *Diário da República n.º 129/2012 – Série I*. Ministério da Educação e Ciência.
- Decreto-Lei n.º 102/2013, de 25 de julho. *Diário da República n.º 142/2013 – Série I*. Ministério da Educação e Ciência.
- Despacho n.º 11838-A/2013, de 11 de setembro. *Diário da República n.º 175/2013 – 1.º suplemento, Série I*. Ministério da Educação e Ciência.

ANEXOS

ANEXO A



INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.
IAVE


U ie Instituto de Educação
LISBOA
UNIVERSIDADE DE LISBOA

Cláusula 1.^a
(Objeto e âmbito)

1. O presente Protocolo tem por objeto a colaboração entre os outorgantes no âmbito da preparação de uma dissertação de Mestrado em Ciências da Educação, especialização em Avaliação Educacional, orientada pelo Professor Doutor Domingos Manuel Barros Fernandes.
2. O projeto referido no número anterior enquadra o trabalho de investigação da Mestranda Maria Antonieta da Costa Mourão Lima Ferreira e pretende analisar a qualidade da adesão das escolas ao Projeto Testes Intermédios, concretamente no que respeita à constituição do elenco de testes a aplicar em cada ano letivo, no triénio 2010-2013.

Cláusula 2.^a
(Obrigações do PRIMEIRO OUTORGANTE)

No âmbito do presente Protocolo, o IAVE compromete-se a disponibilizar ao SEGUNDO OUTORGANTE os dados relativos às inscrições no Projeto Testes Intermédios respeitantes ao 3.º ciclo do ensino básico e ao ensino secundário, nos anos letivos de 2010/2011, 2011/2012 e 2012/2013, com informação sobre as escolas inscritas (com identificação de concelho, área de influência da DRE e tipologia) e sobre os testes intermédios que, em cada um dos últimos três anos, cada escola selecionou para aplicação.

Cláusula 3.^a
(Obrigações do SEGUNDO OUTORGANTE)

O SEGUNDO OUTORGANTE compromete-se a:

- a) Utilizar todos os dados a que tenha acesso, por qualquer forma, no âmbito da execução do presente Protocolo, única e exclusivamente para fins académicos e/ou científicos e no âmbito do projeto identificado na cláusula 1.^a, número 2., estando vedada a sua cedência a terceiros, a qualquer título;

INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.
IAVE



U ie Instituto de Educação
LISBOA
UNIVERSIDADE
DE LISBOA

- b) Sempre que reproduza, cite, ou, por qualquer forma, faça menção aos dados a que tenha acesso no âmbito da execução do presente Protocolo, indicar a fonte e uma caracterização sumária dos mesmos que permita aos destinatários da reprodução, da citação, ou da menção, a sua contextualização com um grau razoável de detalhe, sem prejuízo do sigilo que deva ser observado quanto a referências nominativas ou outras;
- c) Utilizar todos os dados a que tenha acesso, por qualquer forma, no âmbito da execução do presente Protocolo, de forma objetiva, não sendo admissível qualquer forma de tratamento, nomeadamente parcelando a informação, que a desvirtue;
- d) Utilizar todos os dados específicos das escolas a que tenha acesso, por qualquer forma, no âmbito da execução do presente Protocolo, de modo a não ser possível a identificação das escolas a que se referem;
- e) Entregar ao IAVE cópia de todos os estudos, documentos ou trabalhos que sejam objeto de divulgação, por qualquer forma, incluindo publicação, pelo SEGUNDO OUTORGANTE, mesmo que esses estudos, documentos ou trabalhos constituam partes ou rascunhos de estudos, documentos ou trabalhos em elaboração.

Cláusula 4.^a
(Vigência)

O presente Protocolo vigora pelo período que durar a colaboração entre os outorgantes, com a duração máxima de 3 anos, período após o qual o SEGUNDO OUTORGANTE não poderá elaborar, produzir ou divulgar novos estudos, documentos, ou trabalhos que utilizem dados a que tenha tido acesso no âmbito da execução do presente Protocolo, ou que pressuponham o seu conhecimento.

INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.
IAVE

U **ie** Instituto de Educação
LISBOA
UNIVERSIDADE
DE LISBOA

Cláusula 5.^a
(Prorrogação do período de vigência)

O período de vigência supra mencionado poderá ser prorrogado por períodos de 3 anos, por solicitação do SEGUNDO OUTORGANTE.

Feito em duplicado.

Lisboa, 7 de novembro de 2013

O PRIMEIRO OUTORGANTE



O SEGUNDO OUTORGANTE



ANEXO B



GOVERNO DE PORTUGAL

MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA



GABINETE DE AVALIAÇÃO EDUCACIONAL

Projeto Testes Intermédios 2012/2013

3.º Ciclo	Informação n.º 6 PORTUGUÊS - 9.º ANO - 07.fev.2013 Resultados nacionais e regionais
------------------	---

INFORMAÇÃO GERAL

- Número de escolas inscritas: 1177
- Número de escolas que realizaram o teste: 1168 (correspondente a 99,2%)
- Número de escolas que enviaram os resultados: 1158 (correspondente a 98,4%)
- Número de alunos que realizaram o teste: 92 060
- Número de testes resolvidos na íntegra (sem fator de correção): 91 941 (correspondente a 99,9%)

RESULTADOS

Informação geral

Os dados apresentados referem-se apenas aos resultados dos testes resolvidos na íntegra (sem fator de correção).

- Média Nacional: 49,0%
- Desvio Padrão: 17,9%
- Coeficiente de Variação: 36,6%

Distribuição de resultados por item

Grupo I												
	A 1.1.	A 1.2.	A 1.3.	A 1.4.	A 2.	A 3.	B 4.	B 5.	B 6.	B 7.	B 8.	B 9.
% de respostas com cotação máxima	54,91	74,05	55,24	81,04	61,29	14,00	19,51	19,28	11,75	8,90	10,58	15,55
% de respostas com cotação ≥ 75 % e < 100 %						16,72	17,34	17,69	19,66	14,48	13,68	8,66
% de respostas com cotação ≥ 50 % e < 75 %						13,24	10,94	9,15	11,83	8,82	7,99	5,63
% de respostas com cotação ≥ 25 % e < 50 %						17,14	23,93	14,69	20,24	31,92	21,10	19,64
% de respostas com cotação > 0 % e < 25 %						5,07	6,31	2,92	3,64	8,11	4,55	4,22
% de respostas com cotação nula	45,09	25,95	44,76	18,96	38,71	33,83	21,97	36,27	32,88	27,77	42,10	46,30
% da Classificação média/Cotação total	54,91	74,05	55,24	81,04	61,29	43,19	50,78	45,39	43,40	40,17	35,67	34,55

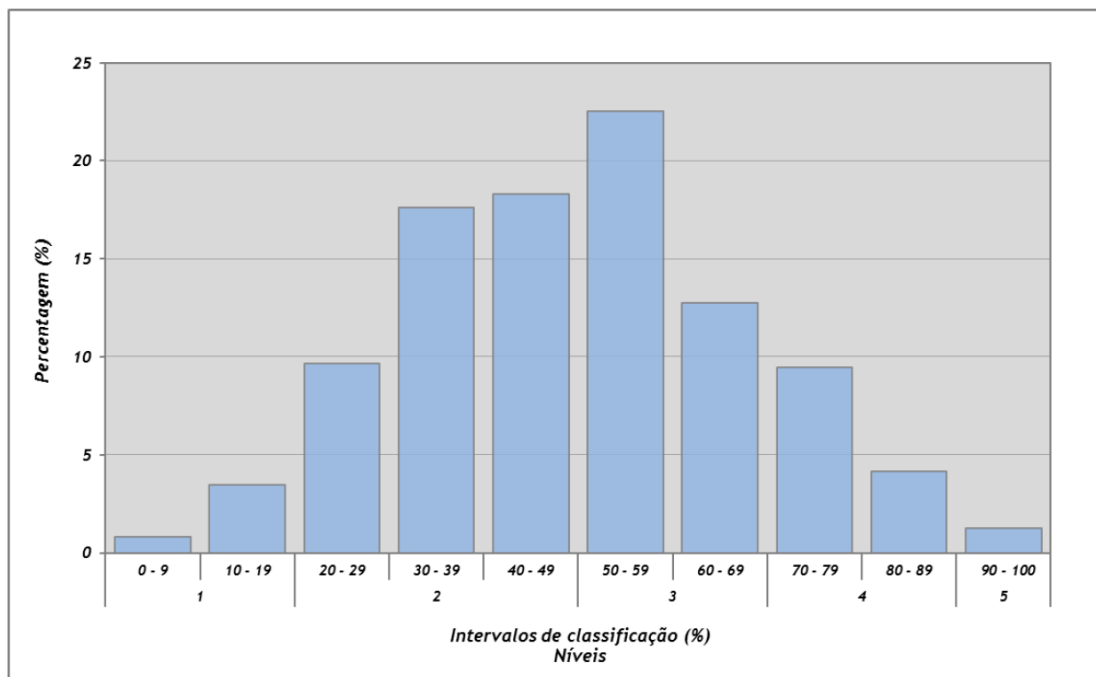
Grupo II

	1.	2.	3.	4.
% de respostas com cotação máxima	18,61	5,77	66,54	3,59
% de respostas com cotação $\geq 75\%$ e $< 100\%$	0,00	4,41		0,00
% de respostas com cotação $\geq 50\%$ e $< 75\%$	31,74	12,77		8,34
% de respostas com cotação $\geq 25\%$ e $< 50\%$	27,05	0,01		25,38
% de respostas com cotação $> 0\%$ e $< 25\%$	15,63	20,00		0,01
% de respostas com cotação nula	6,97	57,04	33,46	62,68
% da Classificação média/Cotação total	51,38	20,96	66,54	17,61

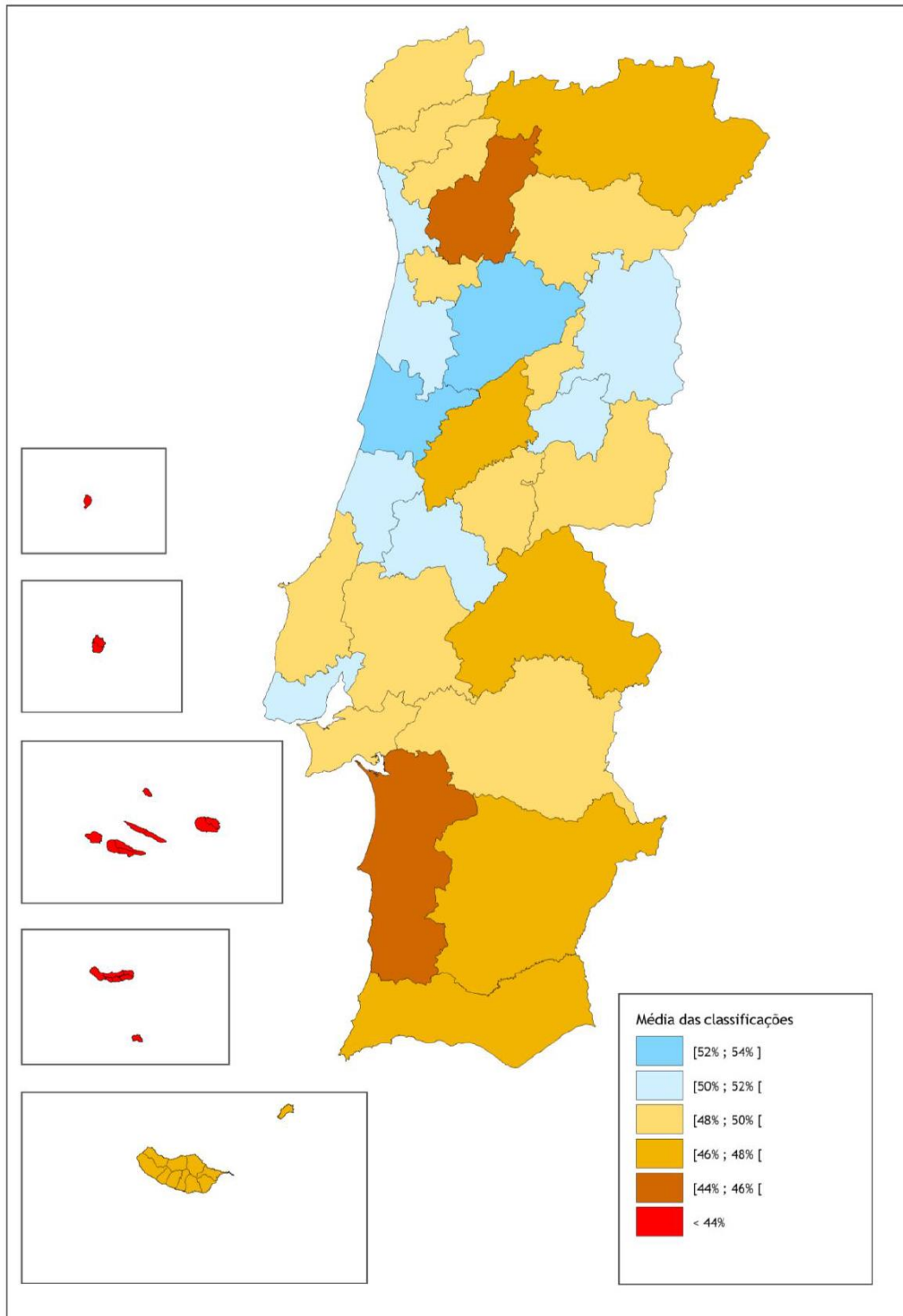
Grupo III (por parâmetro)

	A	B	C	D	E	F	Total
% de respostas com cotação máxima	17,42	10,36	5,32	5,79	6,54	11,18	0,71
% de respostas com cotação $\geq 75\%$ e $< 100\%$	27,16	25,54	20,89	20,46	21,06	30,76	20,58
% de respostas com cotação $\geq 50\%$ e $< 75\%$	29,02	34,70	41,42	42,43	47,58	28,97	52,71
% de respostas com cotação $\geq 25\%$ e $< 50\%$	11,99	16,62	20,68	20,18	14,45	15,48	16,97
% de respostas com cotação $> 0\%$ e $< 25\%$	7,54	5,74	4,67	4,14	3,41	5,86	2,16
% de respostas com cotação nula	6,87	7,04	7,02	7,00	6,96	7,75	6,87
% da Classificação média/Cotação total	62,87	59,40	56,10	56,52	58,40	60,53	57,75

Distribuição das classificações



Distribuição da média das classificações (percentagem) por NUTS III



Publicado em 6.mar.2013

ANEXO C

Elenco de TI disponibilizados (disciplina e ano) 2005-2015, por ciclo de escolaridade (3.º CEB e ES)

	2005/2006	2006/2007	2007/2008	2008/2009	2009/2010	2010/2011	2011/2012	2012/2013	2013/2014	2014/2015	
3.º CEB						CFQ 9	CFQ 9	FQ 9			
						CN 9	CN 9	CN 9			
						Geo 9	Geo 9	Geo 9			
						Hist 9	Hist 9	Hist 9			
						Ing 9	Ing 9	Ing 9	Key for Schools**	Preliminary for Schools**	
					LP 9	LP 9	LP 9	Port 9	Port 9		
			Mat 8	Mat 8	Mat 8	Mat 8	Mat 8				
			Mat 9 t1	Mat 9 t1	Mat 9 t1	Mat 9 t1					
			Mat 9 t2	Mat 9 t2	Mat 9 t2	Mat 9 t2	Mat 9	Mat 9	Mat 9		
	TI - N.º	•	•	3	3	4	9	8	7	2	•
Disc. - N.º	•	•	1	1	2	7	7	7	2	•	
ES			BG 10*	BG 10*	BG 10	BG 10	BG 10				
			BG 11*	BG 11* t1	BG 11 t1	BG 11 t1	BG 11 t1	BG 11	BG 11		
				BG 11* t2	BG 11 t2	BG 11 t2					
						Filos 10					
							Filos 11	Filos 11			
			FQA 10*	FQA 10*	FQA 10	FQA 10	FQA 10				
			FQA 11*	FQA 11* t1	FQA 11 t1	FQA 11 t1	FQA 11	FQA 11	FQA 11		
				FQA 11* t2	FQA 11 t2	FQA 11 t2					
								Ing 11 (nível cont.)	Ing 11 (nível cont.)		
			Mat A 10	Mat A 10 t1	Mat A 10 t1	Mat A 10	Mat A 10				
				Mat A 10 t2	Mat A 10 t2						
	Mat A 11	Mat A 11	Mat A 11 t1	Mat A 11 t1	Mat A 11 t1	Mat A 11 t1	Mat A 11	Mat A 11	Mat A 11		
			Mat A 11 t2	Mat A 11 t2	Mat A 11 t2	Mat A 11 t2					
	Mat A 12 t1	Mat A 12 t1	Mat A 12 t1	Mat A 12 t1	Mat A 12 t1	Mat A 12 t1	Mat A 12	Mat A 12	Mat A 12		
	Mat A 12 t2	Mat A 12 t2	Mat A 12 t2	Mat A 12 t2	Mat A 12 t2	Mat A 12 t2	Mat A 12	Mat A 12	Mat A 12		
				Mat A 12 t3	Mat A 12 t3						
					Mat B 10			Port 12	Port 12	Port 12	
			Mat B 11		Mat B 11						
	TI - N.º	4	3	9	16	15	12	10	7	7	1
	Disc. - N.º	2	1	3	3	4	4	5	5	5	1
TI - N.º Total	4	3	12	16	19	21	18	14	9	1	

* Nos anos letivos assinalados os TI de FQ A e de BG dirigiam-se ao 1.º ou ao 2.º ano de leção da disciplina (que podia ser iniciada no 10.º ou no 11.º ano), não havendo correspondência direta com o ano de escolaridade. ** As provas assinaladas estão contempladas neste histórico apenas porque são referidas nos documentos relativos à apresentação anual do Projeto TI. Correspondem a intervenções ao abrigo do Projeto Key for Schools PORTUGAL (cf. Despacho n.º 11838-A/2013, de 10 de setembro) e não constituem, na sua natureza e finalidades, TI, pelo que não são consideradas para efeitos de caracterização do elenco.

ANEXO D



GOVERNO DE
PORTUGAL

MINISTÉRIO DA EDUCAÇÃO
E CIÊNCIA



INSTITUTO
DE AVALIAÇÃO
EDUCATIVA, I.P.

Projeto Testes Intermédios 2014/2015

INFORMAÇÃO 1

Informação aos Diretores, gestores de projeto e professores

I. Oferta de testes intermédios para 2014/2015

No ano letivo 2014/2015, o IAVE, I.P. irá disponibilizar testes intermédios de Português e de Matemática do 2.º ano de escolaridade. Reitera-se a mais-valia destes testes, porquanto se constituem como instrumentos de regulação externa da aprendizagem nos primeiros anos de escolaridade. A sua aplicação, a meio de um período de escolarização de quatro anos, permite confirmar ou detetar precocemente a existência de problemas de aprendizagem cuja superação, nos anos letivos subseqüentes, é essencial como medida de prevenção do insucesso. A natureza detalhada da informação que os testes geram, bem como a sua adequada utilização pelos professores e pelos encarregados de educação, pode ainda constituir-se como uma ferramenta formativa para auxiliar o desenvolvimento de estratégias específicas de apoio aos alunos que evidenciem mais dificuldades de aprendizagem.

Em relação aos testes aplicados no 3.º ciclo do ensino básico e no ensino secundário, considerando as evidências que os relatórios de anos anteriores têm revelado, nomeadamente o facto de ser residual, ou mesmo impercetível, o seu impacto na melhoria do desempenho dos alunos, entende-se, passados quase 10 anos desde o início do projeto, já não se justificar a sua continuidade nos moldes vigentes.

Atendendo ao vasto espólio de itens disponíveis no Banco de Exames e Provas, considera-se que fica assegurada a possibilidade de estes itens serem usados na familiarização com as provas de avaliação externa ou na autoavaliação dos alunos, sendo a publicitação anual de novos testes uma ação redundante e que pouco acrescenta à informação já existente.

Constitui uma exceção à opção atrás referida a oferta do teste intermédio de Inglês (nível de continuação), a aplicar no 11.º ano de escolaridade, nos moldes adotados no ano letivo anterior. A aplicação deste teste, aferido para o nível B2 do Quadro Europeu Comum de Referência (QEER), enquadra-se numa estratégia de progressiva familiarização de alunos e professores com testes cujo nível de proficiência linguística de referência visa antecipar a prevista adoção, no nosso sistema de ensino, de um teste *First for Schools* (FCE). A aplicação deste teste será feita em contexto similar ao observado em 2014 com a aplicação do teste *Key for Schools* e do que se irá verificar, no corrente ano letivo, com a aplicação do teste *Preliminary for Schools* (PET), no 9.º ano de escolaridade.

Independentemente do elenco de testes a disponibilizar no corrente ano letivo, mantém-se em aberto a possibilidade de, caso a caso, poder vir a ser disponibilizado, transitoriamente, um teste intermédio sempre que seja necessário introduzir numa dada prova final de ciclo ou num exame final nacional alterações profundas na estrutura ou na tipologia de itens a utilizar.

II. Calendário de Aplicação

Ciclo de ensino / Ano de escolaridade / Disciplina			Data de Realização	
			2.º Período	3.º Período
1.º CEB	2.º ano	Português		28 maio
		Matemática		3 junho
ES	11.º ano	Inglês (Continuação)	3 março	

Teste Intermédio de Inglês (nível de continuação) do 11.º ano

- Formação Geral ou Formação Específica
- Compreensão Escrita, Produção e Interação Escritas e Compreensão do Oral – 3 de março
- Produção e Interação Oraís – em momento a definir pela escola, entre 2 de fevereiro e 20 de março

III. Divulgação de informações, período e condições de inscrição

Ano de escolaridade / Disciplina		Datas	
		Divulgação de Informações Projeto e de Informação-teste	Inscrição
2.º ano	Português	Até 16 de março de 2015	7 a 24 de abril de 2015
	Matemática		
11.º ano	Inglês (Continuação)	Até 28 de outubro de 2014	2 a 16 de dezembro de 2014

Não serão aceites inscrições fora do prazo e não serão abertos prazos suplementares para a inscrição.

Complementarmente a esta informação, os períodos de inscrição serão divulgados pela DGEstE, através dos canais habituais, sendo igualmente lembradas as datas referidas na página eletrónica do IAVE, I.P. em www.gave.min-edu.pt e na Extranet em <http://extra.gave.min-edu.pt/login.jsp>.

ANEXO E



GOVERNO DE
PORTUGAL

MINISTÉRIO DA EDUCAÇÃO
E CIÊNCIA

INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.
IAVE

Projeto Testes Intermédios 2013/2014

INFORMAÇÃO-PROJETO

Informação aos Diretores, gestores de projeto e professores

LISTA DE CONTEÚDOS

1. APRESENTAÇÃO E FINALIDADES DO PROJETO	2
2. PRESSUPOSTOS E CONDIÇÕES DE REALIZAÇÃO	4
3. INSCRIÇÃO – PRAZOS E CONDIÇÕES	5
4. DOCUMENTOS DISPONIBILIZADOS E CONDIÇÕES DE ACESSO	6
5. ASPETOS DE ORDEM LOGÍSTICA E ORGANIZACIONAL	7
6. REPRODUÇÃO E DISTRIBUIÇÃO DOS TESTES	8
7. HORÁRIO E CONDIÇÕES DE APLICAÇÃO DOS TESTES	9
8. CORREÇÃO E CLASSIFICAÇÃO DOS TESTES. ANÁLISE DOS RESULTADOS	9
9. PREENCHIMENTO E DEVOLUÇÃO DE GRELHAS DE CLASSIFICAÇÃO E DE QUESTIONÁRIOS DE APLICAÇÃO	10
10. COMUNICAÇÃO COM O IAVE	10
ANEXOS	12
ANEXO A – CALENDÁRIO DAS APLICAÇÕES E DURAÇÃO DOS TESTES	12
ANEXO B – DOCUMENTO SÍNTESE DE PRAZOS DE DEVOLUÇÃO	13
ANEXO C – ADAPTAÇÕES DE TESTES	14



1. APRESENTAÇÃO E FINALIDADES DO PROJETO

No âmbito da missão do Instituto de Avaliação Educativa, I.P. (IAVE), dar-se-á continuidade ao Projeto Testes Intermédios iniciado pelo GAVE no ano letivo de 2005/2006 (nesse ano, apenas para a disciplina de Matemática do 12.º ano). A aplicação dos testes intermédios (TI) teve como primeira finalidade contribuir para a progressiva familiarização de todos os intervenientes com os instrumentos de avaliação sumativa externa, e foi nesse sentido que se estendeu a oferta dos TI a outras disciplinas sujeitas a exame nacional.

O reconhecimento de que os TI podem ter um carácter eminentemente formativo – na medida em que, por um lado, permitem aos alunos a consciencialização da progressão da sua aprendizagem e, por outro lado, permitem aos professores a regulação das suas práticas, tendo por referência padrões de desempenho de âmbito nacional, mediante uma reflexão sustentada pela análise do processo de resposta dos alunos e pelos resultados atingidos – esteve na origem do alargamento da oferta dos TI a disciplinas não sujeitas a prova final de ciclo/exame final nacional.

A avaliação da implementação do projeto em anos anteriores permite concluir que a adesão ao projeto por parte das escolas alcança valores superiores a 95% quando se trata de TI em disciplinas sujeitas a prova final de ciclo/exame final nacional, com exceção do TI de Filosofia, no 11.º ano, cuja adesão não foi além de cerca de 70% das escolas. No que se refere aos TI do 2.º ano (1.º CEB), que visam o diagnóstico precoce das dificuldades dos alunos e uma intervenção pedagógica e didática focada no aumento dos níveis de sucesso no final do 1.º ciclo, a adesão atinge valores da ordem dos 90%.

Em sentido oposto, a adesão aos TI de disciplinas do 3.º CEB não sujeitas a prova final, criados em 2010/2011, como são os casos de Ciências Naturais, Físico-Química, Geografia, História e Inglês, cuja realização deveria ter um enfoque estritamente formativo, orientado para apoiar a regulação externa da aprendizagem e a promoção da elevação da qualidade dos desempenhos dos alunos, mostram níveis de adesão que, em média, não ultrapassam a casa dos 50%, nunca indo além dos 70% das escolas, a que se junta uma tendência decrescente de adesão em todas as disciplinas nos três anos de aplicação.

Relativamente ao TI de Filosofia, disponibilizado desde 2010/2011 com a finalidade de familiarizar os alunos com o exame final nacional facultativo, que viria a ser reintroduzido no elenco dos exames nacionais, os resultados mostram alguma estabilidade, estando os desempenhos mais fracos relacionados com fragilidades de carácter transversal (como sejam o domínio da escrita e da leitura).

No caso particular do Inglês, disciplina em que se reconhece ser essencial assegurar uma consistente melhoria e elevação do nível de proficiência linguística dos alunos,



a adesão das escolas diminuiu (de 69%, em 2010/2011, para 56%, em 2012/2013), no que se refere à realização das componentes escritas e de compreensão do oral do teste (*reading, writing e listening*). Em relação à realização da componente da produção oral (*speaking*), e não obstante ser esta uma das componentes que justifica maior regulação e estarem disponíveis na página do GAVE guiões de apoio à realização dos procedimentos de avaliação deste domínio, a adesão das escolas começou por ser de 12%, em 2010/2011, baixando para 7%, em 2012/2013, o que significa que o número de alunos que realizou esta parte do teste baixou de 2800, em 2010/2011, para 2400, em 2012/2013, de entre os cerca de 90 mil alunos a frequentar o 9.º ano.

A decisão relativa ao conjunto de aplicações a disponibilizar em 2013/2014 (consultar o Anexo A) não traduz, de forma alguma, a desvalorização da dimensão formativa dos TI, que se pretende até reforçar com a antecipação das datas de aplicação em relação ao ano anterior. Ela deve ser entendida como um momento de reavaliação, também por parte das escolas, das condições efetivas de aplicação dos testes, do papel que têm desempenhado no apoio e na promoção do sucesso dos alunos, e dos recursos humanos e financeiros que mobilizam. No *Relatório Nacional dos Testes Intermédios – 2012*, publicado em fevereiro de 2013, apresentava-se uma síntese das razões que foram sendo reportadas pelas escolas (que se prendem, essencialmente, com questões relacionadas com a operacionalização do projeto) e que podem, em parte, justificar o decréscimo de adesão que se tem vindo a verificar.

Tendo em conta que o IAVE tem vindo a disponibilizar um conjunto significativo de informação (desde os enunciados e critérios de classificação de todos os TI aplicados – e demais instrumentos de avaliação externa –, passando pelos resultados por turma, por escola, por região e a nível nacional, até aos relatórios nacionais anuais, quer dos TI quer das provas finais de ciclo e dos exames finais nacionais, é necessário que as escolas avaliem o impacto dessa informação, bem como a eficácia formativa dos TI aplicados, visando-se o reajustamento da oferta em função das necessidades reais dos alunos.

O teste de Inglês do 9.º ano (3.º CEB), tornado obrigatório pelo despacho n.º 11838-A/2013, de 10 de setembro, será aplicado no âmbito do Projeto Key for Schools PORTUGAL, sendo toda a informação sobre o projeto disponibilizada em página eletrónica própria.

A oferta do TI de Inglês no 11.º ano (nível de continuação, tanto na Formação Específica como na Formação Geral) resulta do reconhecimento da necessidade de uma futura certificação dos conhecimentos em língua inglesa, quer para fins pessoais e académicos, quer para fins profissionais, dos alunos do ensino secundário. Este teste visa, principalmente, a aferição do nível de desempenho dos alunos tendo em conta o *Quadro Europeu Comum de Referência para as Línguas Estrangeiras* e a reavaliação das práticas de sala de aula, assumindo uma finalidade eminentemente formativa.

