

Monitorização e *benchmarking* de resultados escolares (exames e provas de aferição) na rede ESCXEL

Rui Santos
CESNOVA, FCSH/NOVA

A Rede de Escolas de Excelência (ESCXEL) é uma parceria iniciada em 2008, entre a Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa, através do seu Centro de Estudos Sociológicos (CESNOVA), e um conjunto de cinco municípios e respetivas escolas públicas. Hoje inclui oito municípios: Amadora, Batalha, Castelo Branco, Constância, Loulé, Mação, Oeiras e Vila de Rei, contando com a adesão voluntária de todos os agrupamentos de escolas e escolas não agrupadas existentes nos concelhos.

Somos uma rede colaborativa multifuncional, com uma grande diversidade de atividades que não cabe aqui resumir. No âmbito deste seminário, ater-me-ei ao processamento pelo CESNOVA dos resultados das avaliações externas dos alunos, expressas nas classificações médias nos exames nacionais de todos os ciclos do Ensino Básico e Secundário e, até recentemente, nas das provas de aferição no termo do 1º e do 2º ciclos do Ensino Básico (classificações externas adiante designadas genericamente por CE). Fazemo-lo através da construção de indicadores homogêneos que permitem:

- **Padronizar**, criando comparabilidade intertemporal.
- **Resumir** e simplificar padrões e tendências.
- **Mapear** desempenhos e, fundamentalmente, tendências de evolução.

O objetivo destes procedimentos e da disseminação dos indicadores processados pelos parceiros é múltiplo. Desde logo, o de acrescentar uma nova camada de inteligibilidade aos resultados, e especialmente à sua evolução, para informar a perceção, a problematização, a discussão e as orientações de ação de professores e diretores, conselhos de escola, e responsáveis políticos e técnicos municipais.

Em segundo lugar, contribuir para uma cultura de colaboração e *benchmarking* entre os atores da rede, através da identificação de agrupamentos de escolas e das escolas não agrupadas (adiante designados genericamente por “escolas”, por comodidade de expressão) que apresentem tendências mais positivas para um dado nível médio de desempenho, apelando à identificação e à disseminação de boas práticas.

Finalmente, contribuir para a cultura de autoavaliação e para a apresentação de resultados às avaliações externas, sejam elas formais por parte da tutela, ou informais por parte da comunidade que servem.

Importa precisar o nosso papel neste processo: a equipa do CESNOVA tem a seu cargo a elaboração dos indicadores, a formação dos utilizadores para a sua leitura e a elaboração dos relatórios que os apresentam e os interpretam resumidamente.

Cabe aos atores educativos e, particularmente, às escolas analisá-los e identificar os contextos e as práticas que lhes subjazem, discutir e partilhar experiências, para gizarem estratégias tendentes à melhoria contínua, que é no fundo o conceito de excelência que designa a rede.

Passo a apresentar em linhas gerais os principais indicadores que construímos, definindo o seu contributo para o acréscimo de inteligibilidade pretendido, segundo as três funções acima definidas.

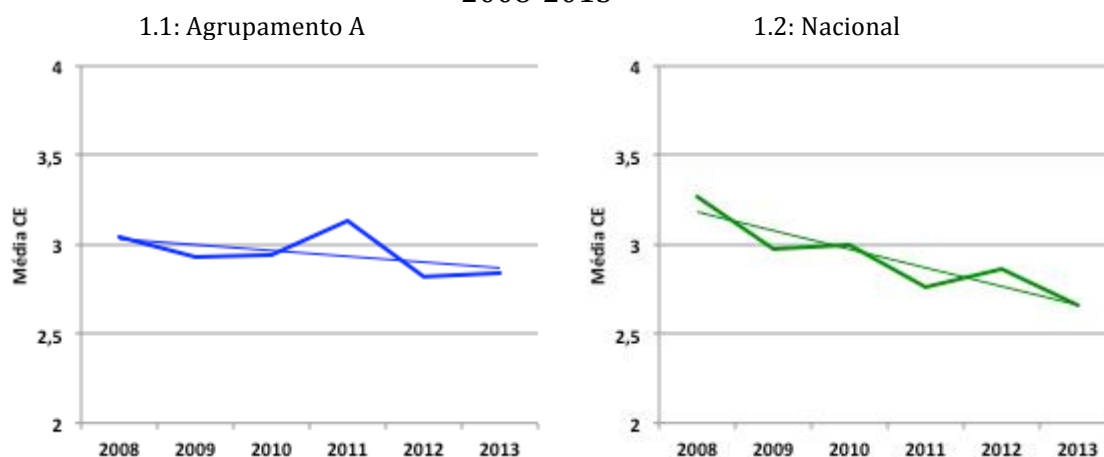
Padronizar

O primeiro princípio em que baseamos a construção dos nossos indicadores é a padronização dos resultados, de modo a que sejam comparáveis entre escolas e municípios e através do tempo, pois só assim podemos inferir conclusões seguras sobre os seus níveis e tendências de evolução.

De facto, cremos que a variabilidade interanual das médias nacionais das CE evidencia um efeito não negligenciável das variações das próprias provas, que pode viciar tanto a apreciação dos níveis médios de resultados nas escolas e municípios num dado ano, como, e sobretudo, a da sua evolução.

Vejamos por exemplo o comportamento recente da média das CE do Agrupamento A da rede ESCXEL nos exames de Português no 9º ciclo (fig. 1.1).

Figura 1: Médias das CE do Agrupamento A e nacional em Português no 9º ano, 2008-2013

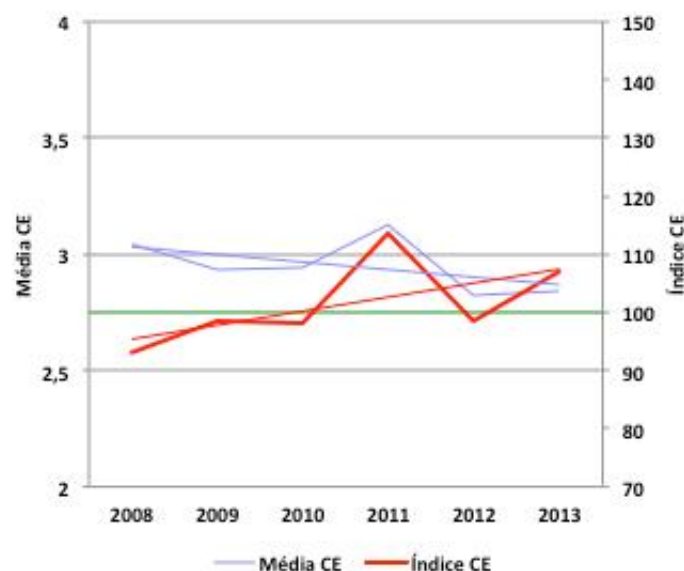


Fonte: Júri Nacional de Exames e ESCXEL, *Agrupamento [A] 2008-2013* (Lisboa: CESNOVA/FCSH, 2014).

A evolução das médias de resultados de exame no agrupamento configura uma evolução negativa ao longo dos seis anos. Contudo, essa evolução ganha um significado claramente diferente uma vez que a confrontamos com a da média nacional da CE, ainda mais negativa (fig. 1.2). Salvo admitindo que as competências dos jovens de todo o país em Português, à saída do 9º ano de escolaridade, tenham simultaneamente oscilado e decrescido ao longo dos seis anos considerados, a evolução explica-se pela variação das provas de exame. Essa é a hipótese que

assumimos. Assim, a evolução do Agrupamento A surge-nos completamente diferente, uma vez que os indexamos à CE média nacional (Fig. 2):

Figura 2: Índice da média da CE do Agrupamento A em Português no 9º ano, base: média nacional da CE=100, 2008-2013



Fonte: Júri Nacional de Exames e ESCXEL, *Agrupamento [A] 2008-2013* (Lisboa: CESNOVA/FCSH, 2014).

Por este motivo, a maioria dos nossos indicadores de resultados escolares baseia-se em índices ou em rácios, tendo como base tornada constante a média nacional no ano de referência (valor 100 do índice CE, na figura 2). Refira-se que, por se tratar de uma indexação à média nacional que anula as diferenças interanuais devidas ao tipo de provas, este tratamento permitiu compatibilizar as novas séries de resultados dos exames nacionais do 4º e do 6º anos do Ensino Básico com as anteriores séries das provas de aferição.

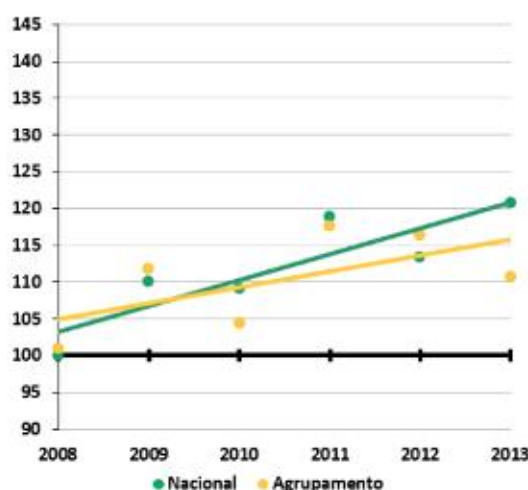
Resumir

Construídos os indicadores, o seu tratamento visa resumir e simplificar a informação em tendências ao longo dos períodos considerados. Nos vários casos em que foram compostos ou recompostos agrupamentos de escolas ao longo de cada período em análise, agregámos retrospectivamente os resultados escolares das unidades (re)agrupadas para toda a janela de seis anos, segundo a composição do agrupamento no último ano. Tal como no exemplo acima, em cada relatório anual trabalhamos um período de seis anos, funcionando numa lógica de janela móvel. Esse intervalo de tempo contém um número de pontos de dados suficiente para identificar uma tendência de evolução, e não tão amplo que essa tendência seja irrelevante para a avaliação das práticas atuais da escola. Esse resumo de tendências é feito à escala da escola e à escala do concelho, quando o concelho tem mais do que um agrupamento.

Relação entre avaliação externa e avaliação interna

Para medir a relação da classificação interna (CIF) com a CE e visualizar a sua evolução, construímos o índice CIF, em que a base 100 constante no tempo é a média CE, à qual indexámos a CIF. Nesta abordagem não assumimos a média nacional como constante, antes comparamos a evolução deste índice em cada disciplina entre a escola e o contexto nacional.

Figura 3: Valores e retas de regressão dos índices CIF do Agrupamento A e nacional em Português no 9º ano, base: médias da CE=100, 2008-2011



Fonte: Júri Nacional de Exames e ESCXEL, *Agrupamento [A] 2008-2013* (Lisboa: CESNOVA/FCSH, 2014).

A figura 3 mostra-nos que, no Agrupamento A, a CIF em Língua Portuguesa no 9º ano esteve sempre acima da CE da disciplina, evoluindo no sentido da valorização relativa da CIF à medida que, como sabemos, as médias da CE desceram (fig. 1.1). À escala nacional, a valorização relativa da CIF foi em geral idêntica, mas a sua subida foi mais acentuada, à medida que as CE desceram também mais acentuadamente, como vimos (fig. 1.2), no fundo sugerindo a inelasticidade da CIF relativamente à CE em ambas as escalas, e maior à escala nacional. Isto mostra às escolas, não só o grau de ajustamento das suas práticas de avaliação interna relativamente aos critérios dos exames nacionais, como em que medida as suas práticas de avaliação interna a aproximam ou diferenciam da tendência nacional de relação entre estas duas classificações.

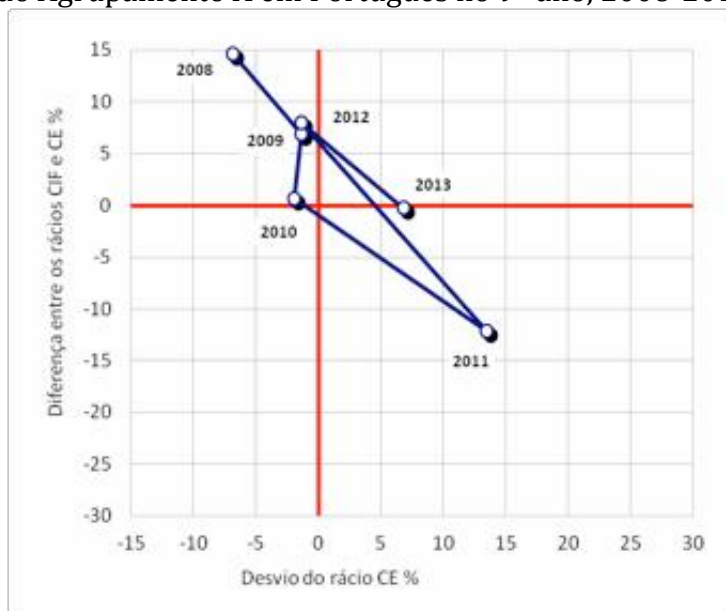
Numa abordagem diversa, resumimos a evolução da relação entre a CE e a CIF operando com dois indicadores de base: o rácio da média CE da unidade de análise (escola ou concelho) à média CE nacional (rácio CE) e o rácio homólogo da média CIF da unidade de análise à média nacional (rácio CIF).

A partir destes, construímos dois indicadores de diferença: o desvio do rácio CE relativamente ao valor de base nacional, que mede em que proporção a média CE da unidade de análise foi superior ou inferior à nacional; e a diferença

entre os rácios CIF e CE, que mede a sub- ou sobrevalorização da CIF relativamente à CE, aferidas pela tendência nacional. A trajetória conjunta destes indicadores num diagrama de fases permite-nos visualizar em que medida a relação entre a sub- ou sobrevalorização da CIF e a CE, aferidas pelos valores nacionais, tendeu, relativamente à relação nacional:

- a permanecer:
 - num espaço de paridade;
 - num espaço de maior exigência da avaliação interna relativamente à externa;
 - num espaço de menor exigência da avaliação interna relativamente à externa;
- a evoluir num sentido ou no outro;
- a ter um comportamento errático.

Figura 4: Relação entre a diferença ao CE e a diferença entre os rácios CIF e CE do Agrupamento A em Português no 9º ano, 2008-2013



Fonte: Júri Nacional de Exames e ESCXEL, *Agrupamento [A] 2008-2013* (Lisboa: CESNOVA/FCSH, 2014).

O exemplo na fig. 4 mostra que o Agrupamento A, na disciplina de Português no 9º ano do Ensino Básico, a relação entre a valorização relativa da CIF e a CE evoluiu entre 2008 e 2011 num sentido de aumento da exigência da CIF, por comparação com o ocorrido à escala nacional, passando ao longo destes quatro anos de uma posição de sobrevalorização relativa da CIF em 2008 para outro de clara sobrevalorização da CE. Nos dois últimos anos do sexénio, regressou a posições mais próximas da relação observada à escala nacional, no triângulo delimitado pelos pontos de 2009, 2010, 2012 e 2013 e para o qual tornou sempre após as posições extremas de 2008 e 2012: valorização relativa da CIF moderada ou nula, combinada com médias CE quase coincidentes, ou moderadamente abaixo das

nacionais. A escola pode assim confrontar esta informação com as suas estratégias de avaliação interna na disciplina, e aferir se esta evolução corresponde aos seus objetivos.

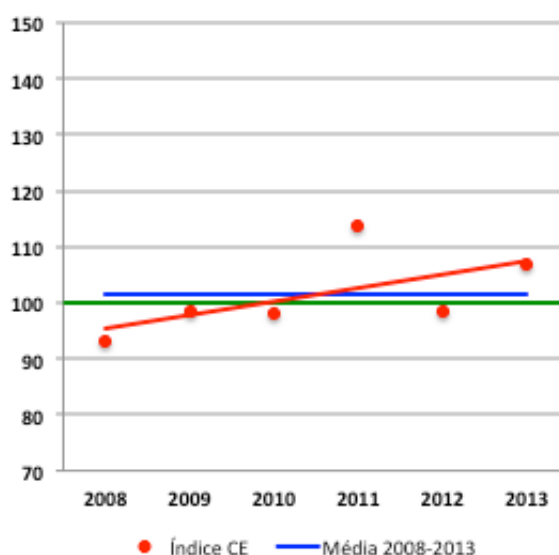
Nível relativo e evolução dos resultados da avaliação externa

Para resumir a evolução dos resultados da avaliação externa, usamos dois indicadores principais:

- o nível médio dos resultados da unidade de análise na avaliação externa, na média das disciplinas e em cada disciplina, relativamente aos resultados nacionais (ou seja, a média do índice CE nos seis anos em análise);
- a evolução tendencial do nível de resultados ao longo do período, dado pelo pendor, positivo ou negativo, e pelo declive mais ou menos acentuado da reta de regressão ajustada aos valores do índice CE ao longo dos anos em análise.

A figura 5 visualiza esses dois indicadores, mostrando que o Agrupamento A teve, no conjunto do período, um nível médio ligeiramente superior ao nacional nos resultados da CE de Língua Portuguesa de 9º ano, e evidenciou ainda uma evolução ascendente, mercê do baixo valor de partida e dos melhores resultados em 2011 e 2013.

Figura 5: Valores anuais, média e reta de regressão do índice CE do Agrupamento A em Português no 9º ano, 2008-2013



Fonte: Júri Nacional de Exames e ESCXEL, *Relatório do Agrupamento [A] 2008-2013* (Lisboa: CESNOVA/FCSH, 2014).

Mapear desempenhos e tendências de evolução

Finalmente, para mapear de modo sintético os desempenhos e as tendências de evolução, utilizamos *scoreboards* (tabelas de pontuações) em que a tendência é

definida pela combinação da média dos índices CE do período considerado com o declive da reta de regressão que resume a sua progressão, segundo as seguintes escalas (fig. 6 e 7):

- atribuímos o *score* 1, "bom", quando a média dos índices CE nos anos considerados é igual ou superior à média nacional da prova (valor 100);
- atribuímos o *score* 2, "sofrível", quando a média é inferior em até 5%, inclusive;
- atribuímos o *score* 3, "em risco", quando a média é inferior em mais de 5% e menos de 15%;
- atribuímos o *score* 4, "mau", quando a média é inferior em 15% ou mais.

Figura 6: *Scores* atribuídos à média, segundo os intervalos de médias dos índices CE

Média	Score
≥ 100	1
$[95, 100[$	2
$]85, 95[$	3
≤ 85	4

O declive, por sua vez, resume a progressão dos resultados, representando a variação linear de a pontos percentuais no índice por cada ano do período observado. No caso de um declive positivo (a superior a 0), a reta é ascendente, o que significa que, no conjunto dos dados observados, a progressão tendencial dos índices CE foi à razão de $a\%$ em cada ano. Inversamente, no caso de um declive negativo (a inferior a 0), a reta é descendente, o que significa que os índices CE tenderam a diminuir $a\%$ em cada ano. No caso de um declive nulo (a igual a 0), a reta é horizontal. Isto significa que, atravessando as variações interanuais, o sentido da progressão dos resultados foi de estagnação ao longo dos anos em análise. Tal como fizemos para a média dos índices CE, atribuímos *scores* a intervalos de valores do declive (fig. 7):

- atribuímos o *score* 1, "bom", quando o declive foi igual ou superior a 2, ou seja, uma progressão tendencial positiva de 2% ou mais por ano;
- atribuímos o *score* 2, "sofrível", quando o declive do Agrupamento foi inferior a 2 e superior a 0, ou seja, uma progressão tendencial positiva, mas fraca;
- atribuímos o *score* 3, "em risco", quando o declive do Agrupamento foi nulo ou negativo mas superior a -2, ou seja, uma progressão tendencial estagnada ou moderadamente negativa;
- estão classificados com o *score* 4, "mau", os casos em que o declive do Agrupamento foi negativo e igual ou inferior a -2, ou seja, uma progressão francamente negativa, com decréscimo tendencial de 2% ou mais por ano.

Figura 7: *Scores* atribuídos ao declive, segundo os intervalos de declives da reta de regressão dos índices CE

Declive	Score
≥ 2	1
$]0,2[$	2
$] -2,0]$	3
≤ -2	4

A tendência, finalmente, resulta da combinação dos dois scores acima. Como mostra a figura 8, supondo que o ritmo de progressão se mantivesse no futuro próximo, consideramos que mesmo uma escola com resultados iguais ou superiores à média nacional estará em risco se tiver tido uma progressão muito negativa (*scores* [1,4]), e uma tendência sofrível se piorou um pouco ou não conseguiu melhorar (1,3); permanecendo com uma boa tendência se tiver melhorado mesmo marginalmente ([1,1] ou [1,2]).

Reciprocamente, uma escola com resultados muito inferiores à média nacional terá, apesar de uma média má, uma tendência sofrível no curto prazo se conseguiu uma evolução francamente positiva (4,1), e menos má, embora ainda em risco, se tiver tido uma progressão positiva sofrível (4,2); mas já se manterá com uma tendência má se não tiver conseguido melhorar nada, ou tiver piorado mesmo que marginalmente ([4,3] ou [4,4]). *Mutatis mutandis*, a mesma lógica aplica-se às situações intermédias.

Figura 8: *Scores* atribuídos à tendência, segundo as combinações dos *scores* da média e do declive

Média	Declive	Tendência
1	1	1
1	2	1
2	1	1
1	3	2
2	2	2
3	1	2
4	1	2
1	4	3
2	3	3
3	2	3
4	2	3
2	4	4
3	3	4
3	4	4
4	3	4
4	4	4

A tabela apresenta os *scores* em sentido descendente, de 1 (verde), os melhores resultados, a 4 (vermelho), os piores. É com estes blocos de cor que

construímos os *scoreboards*, para duas escalas (por escola e por concelho), com diferentes agregações de resultados (por ciclo e por disciplinas em cada ciclo), e representados com diferentes graus de complexidade (desde a média dos índices de todos os exames de só um ciclo até todos os ciclos com as disciplinas discriminadas). A tradução em cores, conforme apresentado nas figuras antecedentes, tornam-nos instrumentos expressivos que permitem uma leitura gráfica rápida dos resultados médios, das suas evoluções e das tendências resultantes no conjunto da rede.

Na ordenação dos *scoreboards* consideramos como primeiro critério as tendências, que resumem as duas variáveis estatísticas de média e declive e, como segundo critério, a média, que corresponde à posição da escola em relação aos resultados médios nacionais dos exames. Esta opção visa acentuar a capacidade de melhorar os resultados, tendo embora em conta o nível médio destes que, como ficou dito, por sua vez resulta em parte da sua evolução.

Tratando-se de instrumentos de leitura simplificados, estes *scoreboards* não dispensam a leitura e a análise dos dados subjacentes; sob os mesmos *scores* de média e declive, podem encontrar-se distribuições substancialmente diferentes de resultados concretos. Por outro lado, pese embora a ordenação, não pretendemos que sejam utilizados como um mero *rankings* de classificação de concelhos ou escolas. Trata-se, sim, de mapas comparativos que resumem o estado dos resultados da rede no cômputo geral do sexénio de referência, em que cada escola pode situar-se e, sobretudo, identificar outras escolas da rede com as quais se possa comparar mais diretamente, tendo em vista o objetivo mais importante: a utilização da rede para *benchmarking*, ou seja, para a aprendizagem partilhada de práticas que a experiência mostre terem resultado bem em situações comparáveis (ou, inversamente, de erros a evitar).

A figura 9 representa o mais recente *scoreboard* da rede relativo ao 9º ano, por escolas, com o período de referência de 2008-2013. Num relance, podemos ver que no conjunto da rede, as escolas apresentam melhores tendências na disciplina de Matemática do que na de Português, e que isso se deve não só a melhores médias (em 25 escolas, 11 tiveram médias boas em Matemática, e apenas nove em Português) mas também, e sobretudo, a melhores progressões dos resultados. Com efeito, seis escolas apresentam bons declives em Matemática, face a apenas uma em Português; disciplina esta em que houve igualmente um caso de declive francamente negativo, ausente na de Matemática. Não menos importante é o facto de só quatro das oito escolas com médias más ou em risco na disciplina de Português apresentarem declives positivos, ainda que sofríveis; enquanto em Matemática, de nove escolas com médias más ou em risco, seis tiveram progressões sofríveis.

Por outro lado, este *scoreboard* mostra que em ambas as disciplinas, as progressões positivas se concentraram nas escolas com médias positivas (12 de 16 em Português e 11 de 17 em Matemática). Em parte, isso deve-se à relação entre o

declive e a média. Como ambos os indicadores foram computados para o mesmo intervalo de tempo, a média não é um valor de partida, mas de chegada. Para situações à partida idênticas, um declive positivo traduz-se num aumento da média, e um negativo numa redução. Logo, se o declive for acentuado, num sentido ou noutro, estará mais provavelmente associado a um melhor ou pior *score* da média. Ainda assim, não deixa de ser preocupante a escassez de declives sofríveis, e a ausência de bons, entre as escolas com médias em risco ou más. Sobretudo atendendo a que, pela relação entre o declive e a média, nas escolas com maus resultados à partida as progressões positivas se traduziriam mais provavelmente em médias de risco ou sofríveis; ora, mesmo entre as escolas com médias sofríveis, só uma (a escola M), e esta apenas em Matemática, apresenta um bom declive de progressão.

Figura 9: *Scoreboard* de Português no 9º ano por agrupamentos, 2008-2013

Concelho	Agrupamento / Escola	PORTUGUÊS			MATEMÁTICA		
		M	D	T	M	D	T
A	A						
B	B						
A	C						
A	D						
B	E						
A	F						
A	G						
C	H						
C	I						
B	J						
D	K						
E	L						
F	M						
A	N						
D	O						
D	P						
G	Q						
A	R						
A	S						
D	T						
D	U						
A	V						
A	W						
B	X						
A	Y						

Fonte: ESCXEL, Relatórios 2008-2013 (Lisboa: CESNOVA/FCSH, 2014).

Numa ótica de *benchmarking*, mais específica a cada escola, poderão por exemplo as escolas W, X e Y, com tendências más nas duas disciplinas, e a V com tendência má em Português, procurar junto das escolas T e U, com médias de risco ou más, identificar o que fizeram para terem alcançado declives positivos em

ambas as disciplinas, assim como a K e a M em Português e as P, Q, R e V em Matemática, etc. O *scoreboard* pode assim servir como um mapa para cada escola identificar possíveis parcerias para aprendizagem dentro da rede, dentro e fora dos seminários que da rede que realizamos regularmente.

Observações finais

Os resultados de exames nacionais não são mais do que uma parte da avaliação externa do trabalho das escolas, por um lado, e da avaliação dos conhecimentos e das competências dos alunos, por outro. Num e noutro terreno, não devem ser tomados como dados absolutos e irredutíveis; as suas determinantes, por outro lado, são certamente complexas e insuficientemente conhecidas. Não podem, contudo, ser ignorados, antes devem ser trabalhados, aprofundada a sua análise e reforçada a sua problematização. O trabalho que prosseguimos na rede ESCXEL e que aqui apresentei em traços largos visa isso mesmo: apresentar de uma outra forma às escolas e aos municípios os dados que, à partida, todos conhecem, processados e organizados de modo a ir além do óbvio, a reduzir a complexidade e, acima de tudo, a estimular a comparação, a interrogação e a colaboração como ferramentas contra a complacência ou o conformismo.