

Comunidades Escolares de Aprendizagem Gulbenkian XXI: exercício metodológico exploratório de análise e estrutura hierárquica de itens e composição sociodemográfica dos perfis de resposta

José Verdasca

jcv@uevora.pt

Centro de Investigação em Educação e Psicologia da Universidade de Évora

Ricardo Monginho

ricardomonjingo@gmail.com

Centro de Investigação em Educação e Psicologia da Universidade de Évora

Ana Cristóvão

anamaria.cristivao@gmail.com

Centro de Investigação em Educação e Psicologia da Universidade de Évora

Rui Espadeiro

ruiespadeiro@gmail.com

Centro de Investigação em Educação e Psicologia da Universidade de Évora

Resumo: No presente artigo, pretendemos, num primeiro momento, evidenciar a importância das práticas colaborativas e participadas na elaboração de instrumentos de avaliação das aprendizagens e da sua aplicação em ambientes virtuais, permitindo, por um lado, acompanhar o desempenho dos alunos em tempo real e, por outro, decorrente da sua transversalidade, induzir processos de autorregulação colaborativa entre professores e microrredes de escolas.

A análise da dificuldade e da discriminação dos itens da prova e das suas interações com recurso ao diagrama de dispersão permitiu projetar as características e qualidade das questões da prova diagnóstica, revelando-se a componente de Português com maior equilíbrio na conjugação dificuldade-discriminação.

O recurso ao algoritmo Exhaustive CHAID para além de permitir apurar que a pontuação global média obtida pelos 133 alunos que realizaram a prova foi de 68,42% e verificar a existência de alguma heterogeneidade dos resultados ($Cv=26,5\%$), permitiu especialmente evidenciar a estrutura hierárquica das respostas e identificar os fluxos sequenciais em termos de resposta certa ou errada na produção da pontuação global. Por outro lado, a análise das sequências estruturais associadas aos nós de maior contraste conduziu ainda ao destaque de seis nós, organizados em torno dos pares de nós (28, 14), (25,18) e (27,24). O primeiro dos pares (28, 14), é também o par

de maior contraste e o nó 28, com maior concentração de respostas certas, o nó com maior peso amostral, representando com os seus 32 casos quase um quarto dos alunos que realizaram a PGD_K3 (24,1%); ao contrário, o nó 14 com apenas 7 casos e representando 5,3% da amostra geral, registou uma pontuação global na prova de 34,2%.

Por último, a análise da estrutura composicional dos perfis em torno de algumas variáveis contextuais sociodemográficas foi reveladora de uma estrutura de desempenho académico na prova que parece continuar mais permeável e refém de fatores contextuais exógenos do que propriamente das dinâmicas internas escolares.

Palavras-chave: análise da dificuldade e discriminação de itens; estrutura hierárquica de itens; estrutura composicional dos perfis.

Learning School Communities Gulbenkian XXI: an exploratory methodological exercise analysis and hierarchical structure of items and sociodemographic composition of response profiles

Abstract: *In this article, we aim at first, highlight the importance of the collaborative and participatory practices in the development of assessment tools of learning and its application in virtual environments, allowing on the one hand, to track student performance in real time and, secondly, due to its pervasive induce autoregulation collaborative processes between teachers and schools' micronetworks.*

The analysis of difficulty and discrimination of the items of evidence and their interactions using the scatterplot allowed to design the characteristics and quality of the questions of diagnostic test, revealing the Portuguese component with greater balance in conjunction difficulty-discrimination. Recourse to Exhaustive CHAID algorithm in addition to allowing the finds that the overall average score for the 133 students who took the test was 68.42% and verify the existence of some heterogeneity of results ($C_v = 26.5\%$), it allowed especially to highlight the hierarchical structure of responses and identify sequential flows in terms of right or wrong answer in the production of the overall score. Furthermore, analysis of structural sequences associated with higher contrast nodes still led to highlight six nodes organized around pairs of nodes (28, 14), (25,18) and (27,24). The first of the pairs (28, 14), is also the pair of higher contrast and the node 28, with the highest concentration of correct responses, the node with larger sample weight, representing with its 32 cases almost a quarter of the students who carried out the PGD_K3 (24.1%); on the contrary, the node 14 with only 7 cases and representing 5.3% of the overall sample, recorded an overall score in the test of 34.2%. Finally, the analysis of the compositional structure of the profiles around some sociodemographic contextual variables was revealing of an academic performance structure in the test that still seems to continue more

permeable and hostage to exogenous contextual factors than properly the school internal dynamics.

Keywords: analysis of difficulty and discrimination items, hierarchical structure of items, compositional structure of the profiles.

Introdução

O projeto *Comunidades Escolares de Aprendizagem Gulbenkian XXI* tem como referência os termos do Programa Gulbenkian Qualificação das Novas Gerações – Promoção de Mudanças na Aprendizagem. Iniciou-se no ano letivo 1415 em três escolas do Alentejo com coortes de alunos do 3º ano de escolaridade (K3) e tem como principal finalidade promover a implementação de um modelo de educação formal baseado numa conceção curricular aberta, cultural e tecnologicamente enriquecida de oportunidades de aprendizagem, pela diversidade de fontes, abordagens, atores, equipamentos.

No âmbito da componente curricular formal, o processo de acompanhamento científico sugere e valoriza entre outros aspetos a promoção de dinâmicas participadas e colaborativas de monitorização, nomeadamente, no âmbito da organização de baterias de questões e de dispositivos de avaliação que permitam informar e disponibilizar na hora os resultados individuais dos alunos e dos grupos de alunos, com ganhos de elevada eficiência em termos de tempo.

No presente artigo pretende-se analisar os resultados da Prova Global Diagnóstica (PGD_K3), aplicada em 8 e 9 de junho às sete turmas do 3º ano de escolaridade envolvidas no projeto PMA-CEAGXXI. A prova é constituída por vinte e oito questões (itens), distribuídas pelas áreas curriculares de Português (sete questões), de Matemática (nove questões) e de Estudo do Meio (doze questões) e foi realizada com recurso ao *Socrative*¹, uma ferramenta multiplataforma que permite criar uma sala virtual e disponibilizar provas aos alunos, bem como acompanhar a sua aprendizagem em tempo real e, se identificados défices, propor e realizar ações imediatas de correção.

A preparação da prova iniciou-se com a construção de uma bolsa de trinta e seis questões, tendo por base os conteúdos curriculares do 3.º ano de escolaridade. Com o

¹ A prova global diagnóstica pode ser acedida e consultada em: https://dl.dropboxusercontent.com/u/23206103/Publica%C3%A7%C3%B5es%20Site%20Projeto/Quiz_provaglobaldediagn%C3%B3sticojunho2015.pdf

objetivo de promover a transversalidade das práticas avaliativas e de induzir processos de autorregulação colaborativa entre professores e microrredes de escolas, partilhou-se um instrumento de avaliação cujas questões foram previamente submetidas à apreciação dos professores titulares das turmas do projeto PMA-CEAGXXI, tendo sido selecionadas vinte e oito das trinta e seis questões inicialmente propostas.

As características da prova dependem, entre outros aspetos, das características psicométricas das suas questões, tendo-se procedido ao cálculo dos parâmetros dos itens no âmbito da Teoria Clássica dos Testes (TCT). As respostas foram tratadas como dicotómicas – certas ou erradas – e comparados os valores obtidos². Num primeiro momento, deter-nos-emos nos índices de dificuldade e de discriminação da prova, que abordaremos de forma conjugada; num segundo momento, procuraremos identificar a estrutura hierárquica das respostas dos alunos aos itens da prova; por último, e em jeito de conclusão, deixaremos algumas reflexões sobre a estrutura composicional dos perfis de desempenho académico com o propósito de mantermos em permanência o debate sobre a responsabilidade sociopolítica e escolar de atores e instituições na construção de dinâmicas curriculares e pedagógicas de promoção da melhoria escolar.

Análise da dificuldade e da discriminação dos itens

O índice de dificuldade de um item está relacionado com a percentagem de respostas erradas dadas a esse item. Na teoria clássica, o índice de dificuldade de um item corresponde à percentagem de alunos que não respondem corretamente a esse item. O seu valor varia entre 0 e 100 pontos percentuais (ou entre 0 e 1), sendo considerado tanto mais difícil quanto maior é o seu índice.

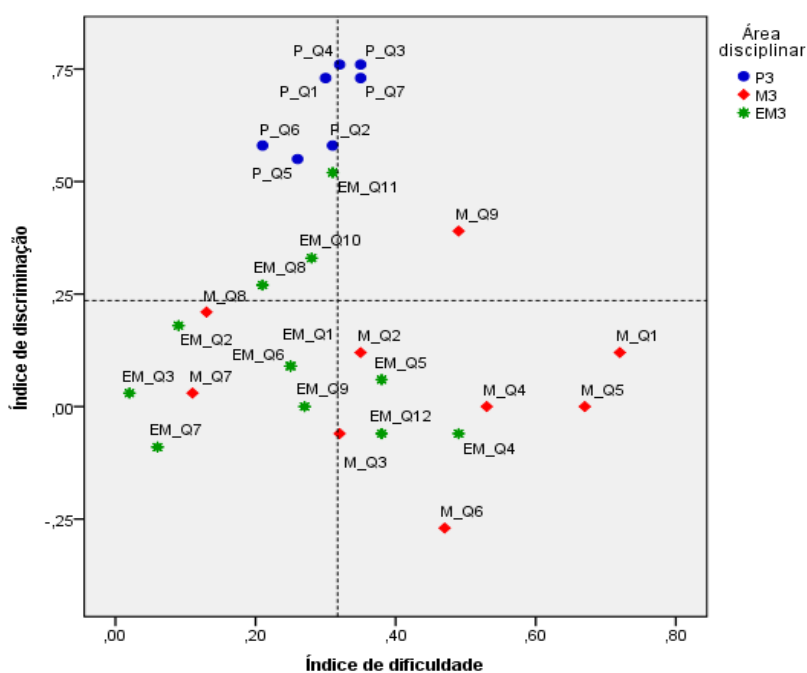
Por sua vez, o índice de discriminação de um item traduz a diferença de sucesso, nesse item, entre os alunos com melhor classificação na prova e os alunos com classificações mais baixas. Os valores de discriminação que se apresentam a seguir traduzem a diferença entre os índices de sucesso dos alunos com classificações mais

² Na presente prova não há lugar a respostas e pontuações politómicas (pontuações intermédias) mas exclusivamente dicotómicas, uma vez que as possibilidades de resposta às diversas questões se apresentavam na forma de escolha múltipla ou de verdadeiro ou falso e gerando, conseqüentemente, respostas absolutamente certas ou erradas e a que se fez corresponder pontuações de um ou zero, respetivamente.

elevadas pelo critério ‘ordem percentual $\geq 75\%$ ’ e os índices de sucesso dos alunos do grupo com classificações mais baixas pelo critério ‘ordem percentual $\leq 25\%$ ’, em cada item da respetiva área curricular. Um índice de discriminação baixo, digamos que igual ou inferior a 0.25, indica que os dois grupos de alunos obtiveram percentagens semelhantes de respostas corretas, no respetivo item. Pelo contrário, índices elevados (igual ou superior a 0.75) revelam diferenças acentuadas entre os índices de sucesso dos alunos melhores classificados na prova e os dos piores classificados.

No diagrama de dispersão projetam-se as 28 questões da prova, através das suas coordenadas referenciadas aos índices de dificuldade e de discriminação, com especificação das respetivas áreas curriculares.

Figura 1. Diagrama de dispersão dos itens da PGD_K3



As linhas tracejadas na vertical e na horizontal representam as respetivas médias dos índices de dificuldade e de discriminação da totalidade das questões da prova e dividem o diagrama em quatro quadrantes. A dificuldade dos itens aumenta da esquerda para a direita e a discriminação de baixo para cima, pelo que, quanto mais as posições relativas dos itens se aproximarem do canto superior direito do diagrama, tanto maior o seu grau de dificuldade e de discriminação e, do canto inferior esquerdo, tanto menor a sua dificuldade e discriminação.

Os resultados mostram:

- um índice médio de dificuldade inferior a 50%, quer na área curricular do Português, quer da Matemática, quer do Estudo do Meio;
- Existem apenas três itens (questões) com um índice de dificuldade superior a 50% (M_Q1, M_Q2 e M_Q3) e todas na área curricular da Matemática;
- O Estudo do Meio é a área curricular que concentra maior número de questões com menor dificuldade (oito em doze) tomando como referência a média global de dificuldade da prova;
- Em termos de discriminação, no quadro do critério definido, apenas na área do Português se observa a existência de duas questões com índices de discriminação superiores a 0,75 (P_Q3 e P_Q4), sendo que, oito dos nove itens de Matemática e oito dos doze de Estudo do Meio, registam índices de discriminação inferiores a 0.25 e, destes, cinco registam mesmo valores negativos de discriminação, significando que, quer os alunos com melhor pontuação na respetiva prova, quer os alunos com piores resultados, obtiveram percentagens semelhantes de respostas corretas.

A observação da posição relativa das 28 questões no diagrama de dispersão sugere a existência de duas ‘nuvens’ posicionais. Uma, concentrando questões nos quadrantes ímpares ou relativamente próximas dos eixos, mostrando a tendência de concentração das questões ao longo de uma trajetória em diagonal e indiciando a existência de uma elevada correlação positiva entre ambos os índices, tendendo a ser os índices de discriminação dos itens da prova tanto mais elevados, quanto maior o respetivo índice de dificuldade. Sobressai ainda da projeção gráfica uma segunda ‘nuvem’ de itens, cujas posições relativas ocupam o quarto quadrante mas em zonas relativamente afastadas dos eixos, sugerindo que, ao contrário das questões da primeira nuvem, a um maior grau de dificuldade dos itens que a compõem não corresponde um maior grau de discriminação. Por outro lado, a aplicação do modelo de regressão pelos mínimos quadrados e o apuramento do respetivo coeficiente de determinação permite confirmar que apenas na área curricular de Português a variação dos índices de discriminação é significativamente explicada pela variação dos índices de dificuldade ($R^2=0,56$), ao contrário das áreas da Matemática e do Estudo do Meio que registam coeficientes de determinação muito baixos.

Em síntese, conclui-se que, em termos de qualidade da prova diagnóstica, a componente de Português regista um bom índice global de qualidade.

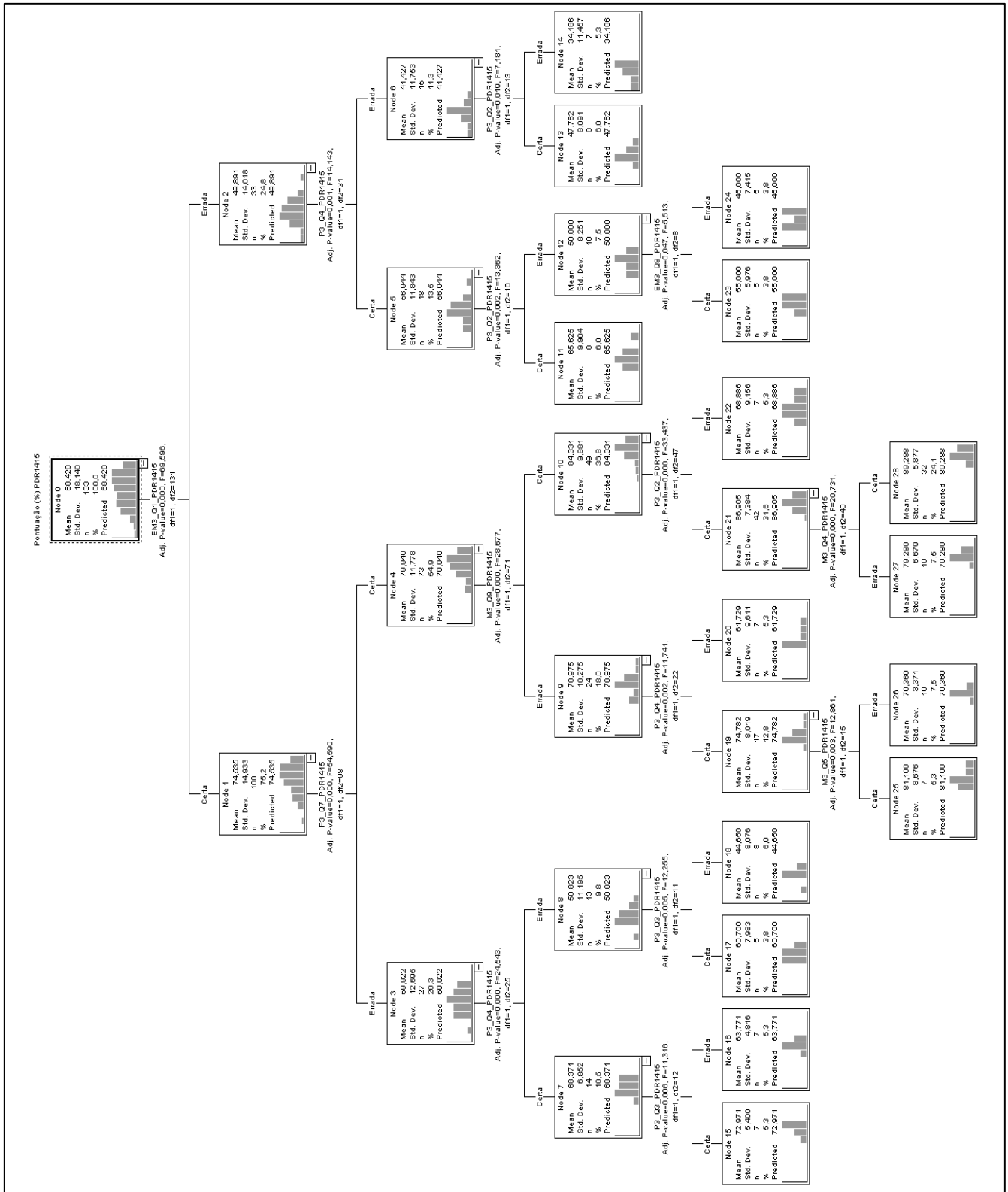
Análise da estrutura das respostas às questões da prova global diagnóstica K3

A identificação dos alinhamentos sequenciais das respostas na produção da pontuação global na prova diagnóstica K3 foi realizada com recurso ao algoritmo EXHAUSTIVE CHAID (*Chi-square Automatic Interaction Detector*), um método estatístico adequado para identificar, classificar e explorar estruturas complexas (Breyman *et al.*, 1984; Pestana e Gageiro, 2009; IBM-SPSS, 2012). Na sua base está a criação de uma estrutura hierárquica – designada de árvore de classificação ou árvore de regressão, consoante a variável de resposta é, respetivamente, quantitativa ou qualitativa, decorrente do fracionamento sucessivo dos dados em subconjuntos mutuamente exclusivos gerando nós intermédios ou nós terminais. A partição dos níveis da árvore faz-se por ordem decrescente de importância das questões. A questão estatisticamente mais significativa define a primeira partição gerando o primeiro nível de profundidade e é avaliada a possibilidade da subdivisão de cada uma das suas categorias de resposta gerando novas segmentações. Este processo é aplicado recursivamente até que a análise esteja terminada. A obtenção de divisões sucessivas dos dados de modo a permitir definir grupos ou classes com base em segmentações significativamente diferentes relativamente à variável dependente constitui o principal objetivo do algoritmo. Qualquer um dos níveis de profundidade da estrutura arbórea pode conter nós terminais, sendo estes a principal base de referência para a identificação das sequências estruturais das respostas dos alunos aos itens da prova.

A solução gerada projetou uma árvore com vinte e nove nós, quinze dos quais terminais, distribuídos por cinco níveis de profundidade (ver Figura 2 e Apêndice 1) e que usaremos na caracterização dos perfis de resposta obtidos nas ramificações sequenciais do nó raiz aos nós terminais da estrutura de classificação gerada.

A estimativa da qualidade do modelo foi obtida mediante validação cruzada, uma alternativa sugerida quando o número de observações é inferior a mil (Pestana e Gageiro, 2009). A proporção de variância explicada pelo modelo é de 85% e, após validação cruzada, de 53% (ver tabela *Risk* em Apêndice 1), sugerindo tais valores que a informação da validação não contraria a aplicação do modelo a outras amostras do mesmo universo.

Figura 2. Estrutura hierárquica das respostas da PGD_K3

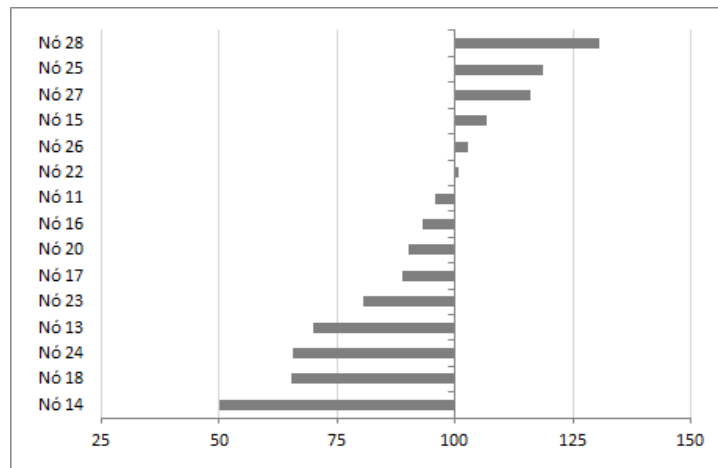


Observando a solução arbórea projetada na figura 2, as respectivas especificações e resultados constantes do Apêndice 1, constata-se que:

- i) As estatísticas do nó raiz (nó 0) mostram que a pontuação global obtida pelos 133 alunos que realizaram a prova foi de 68,42%, com um desvio padrão de 18,14 sugerindo tais valores a existência de alguma disparidade dos resultados obtidos pelos alunos na prova (coeficiente de variação = 26,5%);
- ii) Das 28 questões da prova, face aos critérios de análise estabelecidos, apenas nove e algumas das quais recursivamente, se revelam estatisticamente significativas com uma probabilidade de erro de tipo I (α) inferior a 0,05 na explicação dos perfis sequenciais de resposta e respetiva estrutura hierárquica da prova, pelo que se conclui que nem todas as questões da prova tiveram a mesma importância na configuração final da estrutura hierárquica. Ter acertado ou errado a questão EM_Q₁ não é propriamente o mesmo que ter acertado ou errado a questão EM_Q₂. Enquanto EM_Q₁ emerge como a principal determinante na segmentação da amostra geral em dois grandes ramos, EM_Q₂ nem sequer emerge na solução estrutural obtida com cinco níveis de profundidade. Do mesmo modo, as questões de Português foram mais relevantes no processo de segmentação e desenvolvimento da estrutura do que as questões de Matemática ou de Estudo do Meio, tendo as primeiras interferido em nove segmentações estruturais e as segundas e terceiras apenas em seis.
- iii) É também na componente de Português que se revelam questões recursivas, casos de P_Q₄, P_Q₂ e P_Q₃, determinantes em várias segmentações e em diversos níveis de profundidade e gerando estruturas sequenciais recursivas de respostas.

Apresentam-se de seguida, por ordem decrescente, as médias das pontuações percentuais de cada nó terminal. Para facilidade de visualização dos pares de nós de maior contraste face à amostra geral (Verdasca, 2014), optou-se por projetar graficamente (figura 3) a posição de cada nó relativamente ao nó raiz (base 100) com base nos respetivos *index*, indicativos do grau de concentração de respostas certas dos respetivos nós.

Figura 3. Index dos nós terminais da estrutura hierárquica da PGD_K3



Da projeção gráfica e da tabela que lhe corresponde, sobressaem como pares de maior contraste os pares (28, 14), (25, 18) e (27, 24) com acréscimos de pontuações médias dos primeiros elementos de cada par que superam em 30%, 19% e 16%, respetivamente, o desempenho médio global da amostra geral e com défices de pontuação dos segundos elementos dos respetivos pares de -50%, -35% e -34%. Em termos de diferenciais médios absolutos de pontuações entre os elementos dos pares as diferenças são de 55%, 36% e 34%, respetivamente para o primeiro, segundo e terceiro pares.

A análise da composição e das sequências estruturais associadas a cada um dos nós dos três pares de maior contraste permite constatar que:

- Primeiro par (28, 14)

. *Nó 28*: Para além de ser o nó com maior concentração de respostas certas, é também o nó com maior peso amostral, representando com os seus 32 casos quase um quarto dos alunos que realizaram a PGD_K3 (24,1%) e tendo como perfil de resposta a estrutura sequencial de resposta certa às questões: EM_Q1, P_Q7, M_Q9, P_Q2 e M_Q4, ou seja, [(c) EM_Q1 -> (c) P_Q7 -> (c) M_Q9 -> (c) P_Q2 -> (c) M_Q4];

. *Nó 14*: É o nó com mais baixa concentração de respostas certas. Integra sete alunos que representam 5,3% da amostra, com uma média de pontuações de apenas 34,2%. Como perfil caracterizador deste nó emerge uma estrutura sequencial de respostas erradas aos itens EM_Q1, P_Q4 e P_Q2, ou seja, [(e) EM_Q1 -> (e) P_Q4 -> (e)P_Q2];

- Segundo par (25, 18)

- . *Nó 25*: É o segundo nó com maior concentração de respostas certas, sendo constituído por sete alunos que representam aproximadamente 5% da amostra, tendo como perfil de respostas a seguinte estrutura sequencial: [(c) EM_Q1 -> (c) P_Q7 -> (e) M_Q9 -> (c) P_Q2 -> (c) M_Q4];
- . *Nó 18*: Com a segunda pior média de pontuações (44,65%), os oito alunos que o constituem representam 6% da amostra e a estrutura sequencial de respostas que o caracteriza a seguinte: [(c) EM_Q1 -> (e) P_Q7 -> (e) M_Q4 -> (e) P_Q3];
- Terceiro par (27,24)
 - . *Nó 27*: É o terceiro nó com maior concentração de respostas certas (79,28%) e elevada homogeneidade de resultados (cv=8,4%), sendo constituído por dez alunos com um peso amostral de 7,5% e tendo como perfil de respostas a seguinte estrutura sequencial: [(c) EM_Q1 -> (c) P_Q7 -> (c) M_Q9 -> (c) P_Q2 -> (e) M_Q4];
 - . *Nó 24*: É constituído por 5 alunos, com um peso amostral de apenas 4%. É o terceiro nó com menor concentração de respostas certas (44,65%), com baixa heterogeneidade de pontuações (cv=16%), apresentando a seguinte estrutura sequencial de respostas: [(e) EM_Q1 -> (c) P_Q4 -> (e) P_Q2 -> (e) EM_Q8].

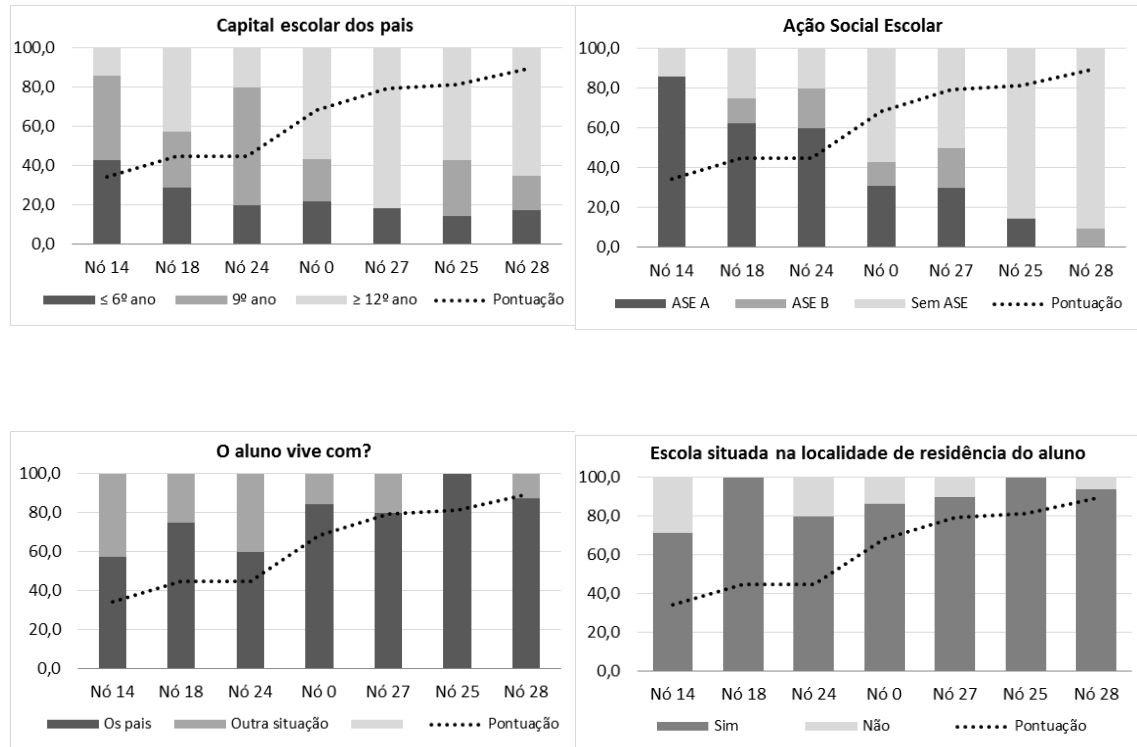
Composição sociodemográfica dos perfis de desempenho

No âmbito dos atributos sociodemográficos dos alunos que realizaram a prova, selecionámos algumas das variáveis comumente descritas na literatura como fortemente condicionantes do desempenho académico dos alunos e das escolas³, caso do capital escolar dos pais, do apoio social escolar, da tipologia familiar do aluno e do local de residência do aluno implicar ou não deslocação diária deste para frequentar a escola. As duas primeiras variáveis são indicativas do contexto cultural e socioeconómico familiar dos alunos; a terceira variável, da vivência no seio de família nuclearizada ou não; a quarta variável, é indicativa de um contexto de habitat de maior ou menor ruralidade. A

³ Referindo-nos apenas a alguns dos estudos e relatórios institucionais mais recentes e de abrangência nacional, veja-se, por exemplo: CNE, 2013, 2014; Canto e Castro *et al*, 2014; Justino *et al*, 2014. No que concerne à turma, enquanto unidade organizacional de análise, a literatura é, desta perspectiva, bem mais escassa e os estudos existentes são, de algum modo, circunscritos a um determinado território (Verdasca, 2002, 2015).

linha tracejada projeta a evolução da pontuação nos sucessivos nós (perfis), sendo que a pontuação média da prova se apresenta registrada no nó raiz (nó 0).

Figura 4. Perfis de desempenho e estrutura composicional sociodemográfica



Mergulhando na configuração contextual de cada um dos nós que integram os três pares de nós de maior contraste, sobressaem claramente linhas tendenciais de diferenciação e segmentação em termos dos níveis de proficiência acadêmica na realização da prova. Esta situação é observável para quaisquer dos fatores contextuais em análise, como se depreende da estrutura de composição de cada um dos perfis (nós) constituídos e das trajetórias ascendentes das pontuações. Os perfis com maiores densidades de alunos cujos pais têm graus de escolaridade mais baixos e maior grau de carência socioeconômica, ou de alunos de famílias não nuclearizadas e com proveniência de habitats de maior ruralidade e isolamento, são também os que tendem a registrar menores índices de desempenho na prova e vice-versa.

Conclusão

No presente artigo, começamos por destacar a importância do uso de dispositivos que em ambiente virtual permitem partilhar e disponibilizar provas aos alunos e acompanhar a sua realização em tempo real, possibilitando ações imediatas de apoio individualizado e correção e, simultaneamente, induzir processos de autorregulação colaborativa entre professores e microrredes de escolas.

Através da observação da distribuição dos respetivos índices de dificuldade e de discriminação das questões da prova no sistema de eixos constatou-se, por um lado, a concentração de um número significativo de questões nos quadrantes ímpares ou relativamente próximas dos eixos, numa trajetória em diagonal claramente indiciadora da existência de correlação positiva elevada entre o índice de dificuldade e o índice de discriminação. Para além desta tendência, no diagrama está projetada uma segunda ‘nuvem’ de itens, cujas posições relativas sugerem que a um maior grau de dificuldade dos itens não corresponde um maior grau de discriminação. Da análise da qualidade das questões da prova diagnóstica, concluiu-se que a componente de Português é que apresenta um maior equilíbrio em termos de dificuldade-discriminação.

A aplicação do método ‘árvores de regressão’ pelo algoritmo Exhaustive CHAID conduziu à projeção de uma árvore com vinte e nove nós, dos quais quinze são nós terminais, e que se apresentam distribuídos por cinco níveis de profundidade, cujas ramificações sequenciais do nó raiz aos nós terminais da estrutura serviram de referência para a caracterização dos perfis de resposta. A análise da composição e das sequências estruturais associadas a cada um dos nós permitiram ainda identificar os nós com maior concentração de respostas certas e de respostas erradas e os respetivos pesos amostrais dos perfis de resposta em termos de estrutura sequencial.

Complementarmente, analisou-se a estrutura composicional dos três pares de nós de maior contraste com base em variáveis contextuais relacionadas com atributos sociodemográficos dos alunos – capital escolar dos pais, escalão da ação social escolar, tipologia familiar e habitat – e compararam-se os resultados obtidos na prova com as diferentes distribuições das modalidades das variáveis em cada um dos nós, sugerindo os resultados apurados que a estrutura de desempenho académico na prova difere significativamente entre os nós segundo as diferentes densidades das modalidades das variáveis. Ou seja: a relação entre a estrutura composicional dos nós de maior contraste e o nível de proficiência académica obtido na prova pelos alunos que compõem esses nós

mostra que os nós (perfis) com maiores densidades de alunos cujos pais têm graus de escolaridade mais elevados e sem carência socioeconómica, ou de alunos de famílias nuclearizadas e com proveniência de habitats urbanos ou tendencialmente urbanos, são também os que tendem a registar maiores índices de desempenho na prova.

Por último, e em jeito de síntese, estamos perante sinais claramente reveladores de uma estrutura de desempenho académico dos alunos que continua a ser marcada por estruturas e lógicas de dominância cultural e social, sintomático de processos escolares aparentemente mais permeáveis e reféns de fatores contextuais exógenos do que das dinâmicas educativas internas (Verdasca, 2015) e que importa fazer notar, tendo em conta a responsabilidade sociopolítica e organizacional da escola (pública) e da comunidade na construção e organização de respostas e práticas curriculares e pedagógicas promotoras da igualdade de oportunidades, equidade e qualidade educativas.

Referências

- Breiman, L., Friedman, J., Olshen, R. e Stone, C. (1984). *Classification and Regression Trees*. Wadsworth International Group. Belmont, CA.
- Castro, L., Santos, J., Pereira, T. e Vitorino, A. (2014). *Modelos para comparação estatística dos resultados académicos em escolas de contexto análogo: Painel de dados para apoio à avaliação externa das escolas*. Lisboa: MEC-DGEEC.
- CNE (2013). *Estado da Educação 2012*. Lisboa: CNE.
- CNE (2014). *Estado da Educação 2013*. Lisboa: CNE.
- IBM SPSS (2012). *Decision Trees 21*. (<ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/21.0/en/client/Mauals/IBMSPSSDecisionTrees.pdf>)
- Justino, D., Pascueiro, L., Franco, L., Santos, R., Almeida, S. e Batista, S. (2014). *Atlas da Educação – Contextos sociais e locais do sucesso e insucesso: Portugal 1991-2012*. Lisboa: CESNOVA.
- Pestana, M. e Gageiro, J. (2009). *Análise Categórica, Árvores de Decisão e Análise de Conteúdo em Ciências Sociais e da Saúde com o SPSS*. Lisboa: Lidel, Edições Técnicas.

Verdasca, J. (2002). *Desempenho escolar, dinâmicas de evolução e elementos configuracionais estruturantes. Os casos do 2.º e 3.º ciclos do ensino básico nos municípios de Évora e Portel*. Évora: Universidade de Évora.

Verdasca, J. (2014). Instrumentos metodológicos e de análise exploratória de dados em educação. (Coletânea de textos e documentos de apoio à disciplina de Administração e Gestão Educacional). Évora: Universidade de Évora (polic.).

Verdasca, J. (2015). A turma como unidade socio-organizacional de análise: contributos para a construção de um roteiro de monitorização e (auto)regulação escolar (Estudos exploratórios e de enquadramento ao modelo 3Q_CDR). CIEP-UEvora (polic.).

Apêndice 1

Prova global diagnóstica K3

Model Summary

Specifications	Growing Method	EXHAUSTIVE CHAID
	Dependent Variable	Pontuação (%) PDR1415
	Independent Variables	M3_Q1_PDR1415, M3_Q2_PDR1415, M3_Q3_PDR1415, M3_Q4_PDR1415, M3_Q5_PDR1415, M3_Q6_PDR1415, M3_Q7_PDR1415, M3_Q8_PDR1415, M3_Q9_PDR1415, P3_Q1_PDR1415, P3_Q2_PDR1415, P3_Q3_PDR1415, P3_Q4_PDR1415, P3_Q5_PDR1415, P3_Q6_PDR1415, P3_Q7_PDR1415, EM3_Q1_PDR1415, EM3_Q2_PDR1415, EM3_Q3_PDR1415, EM3_Q4_PDR1415, EM3_Q5_PDR1415, EM3_Q6_PDR1415, EM3_Q7_PDR1415, EM3_Q8_PDR1415, EM3_Q9_PDR1415, EM3_Q10_PDR1415, EM3_Q11_PDR1415, EM3_Q12_PDR1415
	Validation	Cross Validation
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	10
	Minimum Cases in Child Node	5
Results	Independent Variables Included	EM3_Q1_PDR1415, P3_Q7_PDR1415, P3_Q4_PDR1415, P3_Q3_PDR1415, M3_Q9_PDR1415, M3_Q5_PDR1415, P3_Q2_PDR1415, M3_Q4_PDR1415, EM3_Q8_PDR1415
	Number of Nodes	29
	Number of Terminal Nodes	15
	Depth	5

Gain Summary for Nodes

Node	N	Percent	Mean
28	32	24,1%	89,288
25	7	5,3%	81,100
27	10	7,5%	79,280
15	7	5,3%	72,971
26	10	7,5%	70,360
22	7	5,3%	68,886
11	8	6,0%	65,625
16	7	5,3%	63,771
20	7	5,3%	61,729
17	5	3,8%	60,700
23	5	3,8%	55,000
13	8	6,0%	47,762
24	5	3,8%	45,000
18	8	6,0%	44,650
14	7	5,3%	34,186

Growing Method: EXHAUSTIVE CHAID
 Dependent Variable: Pontuação (%) PDR1415

Risk

Method	Estimate	Std. Error
Resubstitution	48,151	6,396
Cross-Validation	153,697	23,404

Growing Method: EXHAUSTIVE CHAID
 Dependent Variable: Pontuação (%) PDR1415