

## Parecer sobre a prova 715

### Cotações e existência de itens cuja contabilização depende da pontuação obtida

#### Resumo

Este parecer incide sobre dois aspetos da prova 715: a distribuição das cotações pelos itens da prova e a existência de itens cuja contabilização para a classificação final depende da pontuação obtida nesse conjunto de itens. Para a melhoria da validade e fiabilidade da prova, propõe-se que as cotações devam ser diferenciadas de acordo com a tipologia e formato dos itens e que todos os itens a que o aluno responda possam contabilizar efetivamente para a classificação final. No sentido de permitir acomodar diferentes opções de gestão curricular, poderão ser criados grupos de itens em opção.

#### 1. Distribuição das cotações

Dado o duplo papel das provas de exame nacional, a certificação do ensino secundário e o acesso ao ensino superior, estas provas devem, por um lado, ser equilibradas na valorização das várias competências avaliadas e dos níveis de complexidade cognitiva requeridos e, por outro lado, constituírem-se como um instrumento que permita avaliar com equidade o desempenho demonstrado pelos alunos nas respostas dadas aos itens.

A opção de atribuir a mesma cotação (ou muito semelhantes) a itens de diferentes tipologias, independentemente do tempo médio esperado de resposta associado a cada um deles (itens cujas respostas são obtidas em 1 ou 2 minutos têm a mesma cotação de itens cujas respostas podem necessitar de mais de 15 minutos), não garante nem o equilíbrio na valorização das competências avaliadas nem a equidade na classificação atribuída ao desempenho demonstrado nas respostas dadas (examinandos com domínios de conhecimentos e competências muito diferentes podem ter classificações praticamente iguais).

Na prática, verifica-se que a cotação atribuída a alguns dos itens de resposta restrita não permite discriminar o desempenho dos alunos, uma vez que alunos com desempenhos muito

diferentes podem ter a mesma pontuação (por exemplo, nos itens de resposta restrita de cálculo, um aluno que cumpra todas as etapas com dois erros de tipo 2 pode obter uma pontuação igual - desvalorização de 40% da cotação do item -, ou até inferior, a um aluno que apenas cumpriu 50% das etapas a que teria que dar resposta).

Assim, propõe-se a existência de três cotações diferenciadas de acordo com os formatos dos itens:

- menor cotação para itens de construção de resposta curta que apelem a reconhecer e a reproduzir (por exemplo, o nome de uma grandeza ou de uma unidade, a leitura numa escala, a fórmula de estrutura de um composto simples, etc.);
- cotação intermédia para itens de seleção (escolha múltipla, associação, ordenação, verdadeiro/falso e completamento), sendo a classificação dicotómica na escolha múltipla, mas podendo existir níveis de desempenho nos restantes formatos;
- maior cotação para itens de resposta restrita (a resposta implica, por exemplo, a apresentação de uma explicação, de uma previsão, de uma conclusão, de uma justificação, de cálculos ou de determinações gráficas); estes itens podem ser classificados por etapas ou por níveis de desempenho.

Para evitar uma sobrevalorização das competências de maior nível de complexidade, o rácio das cotações dos itens da prova (quociente entre a cotação mais alta e a cotação mais baixa) deve situar-se entre 1,5 e 2.

Nesta proposta, itens de diferentes níveis de complexidade cognitiva podem ter a mesma cotação, mas garante-se não só que possa existir uma melhor adequação entre a diversidade de respostas dadas a um mesmo item e as correspondentes pontuações (maior equidade no processo de classificação e, previsivelmente, maior fiabilidade da prova), como também entre o intervalo de tempos médios de resposta de itens de diferentes formatos e a correspondente cotação (um item de resposta curta que avalia conhecimentos declarativos exige um muito menor tempo de resolução do que um item de resposta restrita em que o examinando tem que estabelecer uma estratégia de resolução com várias etapas).

O equilíbrio nas competências avaliadas deverá ser garantido através de uma distribuição adequada da cotação global associada aos diversos níveis de complexidade cognitiva, por exemplo, 30% para o nível inferior (reconhecer, reproduzir, ...), 40% para o nível médio (interpretar, aplicar, ...) e 30% para o nível superior (raciocinar, fundamentar, ...).

## 2. Itens cuja contabilização para a classificação final depende da pontuação obtida

Desde 2020 que as provas de exame nacional apresentam itens cuja resposta é obrigatoriamente contabilizada para a classificação final e itens cuja contabilização para a classificação final depende da pontuação obtida nesse conjunto de itens (no caso da prova 715, desde 2021 que há um conjunto de 8 itens, todos de seleção, dos quais são considerados, para a classificação final da prova, os 4 itens cujas respostas obtenham melhor pontuação). Esta opção tem sido fundamentada pelo IAVE com base na acomodação de diferentes opções de gestão curricular.

Sendo o Perfil dos Alunos à Saída da Escolaridade Obrigatória e as Aprendizagens Essenciais da disciplina os documentos de referência na conceção das provas de avaliação externa, não é congruente admitir-se que as provas incluam itens que não avaliem aprendizagens essenciais. Do ponto de vista das aprendizagens avaliadas, não existe fundamento para a diferenciação entre os itens cuja resposta é obrigatoriamente contabilizada e os itens em que não há essa obrigatoriedade, uma vez que ambos avaliam aprendizagens significativas e essenciais. Verifica-se que em provas anteriores, itens com aprendizagens e conteúdos que numas provas contabilizaram obrigatoriamente para a classificação final foram noutras provas incluídas nos que poderiam, ou não, contabilizar para essa mesma classificação.

Na verdade, não há neste modelo nenhuma opção de gestão curricular, uma vez que os examinandos podem (e devem) responder a todos os itens. O facto de os examinandos poderem responder, ou não, a todos os itens acarreta problemas de gestão do tempo e pode contribuir para o desajustamento entre a extensão da prova e o tempo disponível para a sua resolução, sem contar com o período de tolerância.

O facto de a classificação depender da escolha das melhores respostas origina diversas injustiças:

- para os examinandos que acertam todos os itens cuja contabilização para a classificação final depende da pontuação obtida, ficam indeterminados os itens que realmente contabilizam para a classificação final;
- classificações iguais podem corresponder a desempenhos muito diferentes (por exemplo, na prova de exame 715, examinandos que acertem X itens e (X+4) itens podem ter a mesma classificação, para X maior ou igual que 4);
- classificações mais baixas podem corresponder a melhores desempenhos (por exemplo, na prova de exame 715, um examinando pode acertar 20 itens e ser classificado com

160 pontos, por ter falhado 4 itens “obrigatórios”, e outro examinado que acerte 18 itens pode ser classificado com 180 pontos, por ter falhado apenas 2 itens “obrigatórios”);

- pode ser mais penalizador para os examinandos que, por exemplo, não lecionaram parte do programa, dado que esses examinandos não podem usufruir da escolha das melhores respostas, por não estarem em condições de responderem a todos os itens.

Assim, considera-se que este modelo viola um dos princípios que deveria estar na base de uma avaliação justa e séria, a correspondência entre o desempenho do examinando e a respetiva classificação. Esta ausência de correspondência assume especial gravidade numa prova que tem também a função de seriação dos alunos no acesso ao ensino superior.

O facto de os examinandos poderem errar e não serem penalizados e a eventual diminuição de ansiedade que essa possibilidade de erro lhes pode dar têm contribuído para que o modelo descrito, adotado pelo IAVE, seja aceite pela comunidade educativa, esquecendo-se da injustiça subjacente ao facto de alguns erros serem penalizados e outros não.

Considera-se que um modelo em que fossem verdadeiramente valorizadas opções de gestão curricular deveria dar a possibilidade aos examinados de escolher entre dois grupos de itens, de níveis de complexidade cognitiva idênticos entre si, versando diferentes conteúdos curriculares.

24 de outubro de 2024

A Direção da Associação Portuguesa de Professores de Física e de Química