

FACTORES RELEVANTES NA AVALIAÇÃO ESCOLAR POR PERGUNTAS DE ESCOLHA MÚLTIPLA ¹

Amâncio da Costa Pinto
Faculdade de Psicologia e de C. da Educação, Universidade do Porto

Resumo

O uso de perguntas de escolha múltipla para avaliar o conhecimento escolar ou outros conhecimentos de índole geral é cada vez mais frequente. Grande parte desta popularidade advém da objectividade e facilidade de correcção das respostas dadas. Todavia este método de avaliação tem algumas limitações, a mais importante das quais é a dificuldade de elaborar um formato de avaliação que seja representativo dos conhecimentos e competências escolares envolvidos. Este artigo aborda um conjunto de questões relacionadas com o uso de perguntas de escolha múltipla em exames, refere as vantagens e inconvenientes em relação a formatos de avaliação escolar por perguntas de desenvolvimento e exemplifica estas limitações a partir das perguntas dos exames nacionais de psicologia do ensino secundário no ano 2000.

Palavras Chave: *Avaliação, perguntas de escolha múltipla, psicologia, exames nacionais, secundário.*

O leitor andaria de avião se soubesse que o piloto tinha sido seleccionado e avaliado no curso de pilotos por meio de perguntas de escolha múltipla (PEMs)? E no caso de um motorista de autocarro, aceitaria entrar no autocarro? E se fosse um médico, concordaria ser submetido a uma intervenção cirúrgica? Quando a pergunta é posta nestes termos, a tendência imediata das pessoas é dizer não, e no entanto o mais provável é que os conhecimentos teóricos em cada uma destas diferentes áreas tenham sido avaliados por meio de PEMs, como aliás aconteceu com cada um de nós quando realizou o exame do código de condução. Naturalmente que o saber do piloto, do motorista e do médico envolve uma conjugação elevada de conhecimentos teóricos e práticos, isto é, por um lado saber, por outro competências e habilidades, ou na linguagem da neuropsicologia e dos psicólogos cognitivos, uma conjugação de conhecimentos *declarativos* e de conhecimentos *procedimentais* (Cohen e Squire, 1980; Anderson, 1990; Pinto, 2001a).

Cada um destes saberes tem características próprias, processos de aquisição, retenção e recordação diferentes e são avaliados também de modo diferente. Os conhecimentos teóricos e os saberes de natureza declarativa podem ser avaliados por meio de perguntas de escolha múltipla (PEMs) ou através de perguntas de completação e de desenvolvimento, enquanto que os saberes da ordem das competências, habilidades, ou do saber fazer (know how em inglês) só podem ser avaliados cabalmente através da observação visual do executante, como pilotar um avião, conduzir um carro ou efectuar uma intervenção cirúrgica (ou através da observação auditiva no caso da música; observação olfactiva no caso de discriminação de aromas complexos, etc.).

¹ Publicação: [Pinto, A. C. (2001). Factores relevantes na avaliação escolar por perguntas de escolha múltipla. *Psicologia, Educação e Cultura*, 5 (1), 23-44].

Morada (address): Faculdade de Psicologia, Universidade do Porto, R. Campo Alegre, 1055, 4169-004 Porto, Portugal. E.mail: amancio@psi.up.pt. Estudo realizado no âmbito do projecto FCT nº 113/94. Partes deste artigo foram apresentados nas V Jornadas Psicopedagógicas de Gaia em 29 de Novembro de 2000.

No sistema escolar há conhecimentos adquiridos muito importantes que são manifestamente da ordem das competências, habilidades e do saber fazer, como a leitura e a escrita, mas há também muitos outros que são de natureza declarativa e podem ser avaliados através de PEMs ou de perguntas de desenvolvimento. A avaliação do conhecimento escolar, ou outros conhecimentos de índole geral, por meio de PEMs é cada vez mais frequente e popular entre os docentes, alunos e instituições. Grande parte desta popularidade entre os docentes advém da facilidade de correção, para os alunos da eliminação da subjectividade de correção reduzindo substancialmente o poder discricionário dos docentes, para as instituições pelos ganhos económicos substanciais que a sua aplicação apresenta quando está em causa a avaliação de milhares de alunos.

O objectivo deste artigo é descrever e analisar mais atentamente as vantagens e inconvenientes que a avaliação por PEMs apresenta relativamente à avaliação alternativa por perguntas de desenvolvimento.

A avaliação por PEMs é complexa e as vantagens da sua aplicação nem sempre são tão substanciais que levem um docente a adoptá-la precipitadamente. Mesmo os professores que têm alguma simpatia por exames constituídos por PEMs não adoptam este sistema de forma incondicional, preferindo sistemas mistos de PEMs e de perguntas de desenvolvimento. Eu próprio usei a avaliação por PEMs em exclusivo numa cadeira universitária de que sou regente, de início com grande entusiasmo, para mais tarde, passados cinco anos, abandonar este sistema de avaliação. Reconheço que as razões do abandono foram mais marginais do que substanciais, mas mesmo que venha a retomar a avaliação por PEMs no futuro, o entusiasmo nunca será o mesmo como no primeiro ano de aplicação.

Perguntas de Escolha Múltipla e Perguntas de Desenvolvimento

As perguntas de escolha múltipla (PEMs) e as perguntas de desenvolvimento (PDs) fazem parte da maioria dos exames a que um estudante é submetido ao longo do processo escolar (vide exemplos na II Parte). As PEMs são habitualmente designadas por provas “objectivas”, embora o elemento objectivo destas provas se limite apenas ao sistema de correção que é bastante mais fiel do que o seguido por sistemas alternativos e que pode até ser feito de forma mecânica. As provas objectivas de avaliação são formadas por vários tipos, onde se incluem entre os tipos mais frequentes: (1) perguntas de escolha múltipla com 3 a 5 opções, às vezes mais; (2) perguntas de escolha múltipla emparelhada (Vide II Parte, pergunta 1); (3) perguntas de escolha dupla verdadeiro-falso; (4) perguntas de completação de palavras. As provas objectivas de avaliação mais frequentemente usadas adoptam o formato de perguntas de escolha múltipla com quatro opções (PEMs-4). Num exame escolar anual constituído integralmente por PEMs-4, o número de questões para avaliar o conhecimento escolar é de 40 a 50 para uma prova com uma duração de 60 a 90 minutos.

As provas formadas por perguntas de desenvolvimento (PDs) são constituídas por um enunciado, sob a forma de pergunta ou afirmação, onde o aluno é solicitado a responder por escrito evocando conhecimentos, e procedendo em maior ou menor grau a uma análise, descrição, explicação, comentário e síntese de um conteúdo, tópico ou área do saber escolar. Cada um destes termos tem significados próprios, que deve ser tido em conta na elaboração da pergunta de exame (Vide II Parte, pergunta 11). A resposta por escrito pode ser mais ou menos extensa e varia de algumas frases até vários parágrafos. Num exame escolar constituído integralmente por PDs, o número de questões para avaliar o

conhecimento escolar é normalmente de 5 a 7 para uma prova com uma duração média de 120 minutos.

Inserir o Quadro 1 ± aqui

Um exame constituído integralmente por PEMs ou por PDs tem características específicas, cuja descrição e análise apresento no Quadro 1 ao longo de 24 elementos comparativos. A maior parte dos factores comparativos deste Quadro 1 são análises pessoais, que têm por base uma comparação bastante mais elementar efectuada por Gronlund (1988). Estas 24 comparações entre PEMs e PDs podem agrupar-se em torno de quatro categorias: (1) Aspectos organizacionais do exame, constituído pelos factores de 1 a 7; (2) Correção das questões, constituído pelos factores de 8 a 13; (3) Preferências e processos cognitivos envolvidos, constituído pelos factores de 14 a 21; (4) Aspectos vários, constituído pelos factores de 22 a 24. Muitas das diferenças entre PEMs e PDs expressas no Quadro 1 são auto-explicativas para um docente com alguma experiência neste tipo de provas, ficando assim dispensado de as caracterizar mais desenvolvidamente por razões de espaço, com excepção dos factores de 14 a 17 que requerem uma análise mais desenvolvida.

Provas de memória de reconhecimento e de evocação (#14)

A avaliação por PEMs está associado à provas de memória por *reconhecimento* e a avaliação por PDs a provas de memória por *evocação*. Numa PEM-4 a resposta correcta está presente numa das 4 alternativas e o estudante deve apenas *reconhecê-la* e assinalá-la com uma cruz ou escrever a letra que a identifica numa folha de respostas. Numa PD o aluno tem primeiro de *evocar* e recordar as informações correctas, depois organizá-las mentalmente num formato de resposta que pareça adequado e depois escrever a resposta numa escrita legível, agradável, gramaticalmente correcta e sem erros de ortografia.

Nas PDs o número de competências cognitivas envolvidas é maior e a prova de memória por evocação é também mais complexa e difícil. De facto nas tarefas do dia a dia reconhecemos mais facilmente a familiaridade de um rosto e de uma voz do que a evocação ou a recordação do nome da pessoa. No laboratório, o desempenho médio de memória em provas de evocação é da ordem dos 40 a 60%, bastante inferior ao desempenho em provas de reconhecimento (Baddeley, 1997; Pinto, 1992, cap. 12 e 13; Pinto, 1998).

Um prova por PEM é aparentemente uma tarefa simples, porque muitas das tarefas no dia a dia a que estamos sujeitos nos parecem de reconhecimento simples, como acontece com o reconhecimento de pessoas, seres, objectos e acontecimentos cuja familiaridade ou falta dela imediatamente detectamos. No laboratório, o desempenho médio de memória em provas de reconhecimento é bastante mais elevado do que nas provas de evocação e situa-se na ordem dos 70 a 80%. Por estas ou outras razões, os estudantes iludem-se frequentemente sobre a facilidade das provas por PEMs e preferem-nas em relação a provas por PDs. No entanto as provas por PEMs, apesar de envolverem processos de reconhecimento aparentemente mais simples, podem tornar-se bastante complexas e difíceis, quando o número de alternativas é maior e os distractores são bastante similares entre si.

Categorias de pensamento (#15)

Será possível planear uma avaliação escolar de modo que o resultado reflecta algo mais do que uma simples memória do que está escrito no manual ou foi referido na aula? Pode-se ir além da memória do que se aprendeu? Para se responder é preciso conhecer quais os processos mais elevados de pensamento, assim como quais os métodos de medição específicos destes processos que podem ser avaliados no meio escolar. Requer-se assim uma tipologia dos processos de pensar. Nas últimas décadas os investigadores propuseram várias tipologias (e.g., Bloom et al., 1956; Gagné, 1968; Miller et al. 1978; Royer et al., 1993). Destas tipologias, a tipologia de Benjamin Bloom é uma das mais antigas, conhecidas e estudadas. Bloom sugeriu seis níveis de avaliação de conteúdos escolares a nível cognitivo que poderiam ser representativos de processos de pensamento mais elevados.² Os objectivos educativos de ordem cognitiva são os seguintes:

1. *Conhecimento*: Recordação e evocação de informação previamente aprendida em termos de factos, conceitos, princípios, teorias e procedimentos. O objectivo mais baixo na hierarquia de Bloom. Exemplos: Indicar nomes, autores, datas, etc.
2. *Compreensão*: Capacidade para captar o significado da informação e saber interpretá-la e traduzi-la em várias modalidades. Exemplos: Referir elementos de uma categoria; tradução de uma frase de uma língua para outra; inferir o resultado de uma história; interpretar um quadro ou gráfico.
3. *Aplicação*: Capacidade para usar a informação escolar num novo contexto, empregando conceitos e princípios em situações novas. Exemplos: Usar um barómetro para prever o tempo; aplicar princípios de aprendizagem na organização do estudo de uma disciplina.
4. *Análise*: Capacidade para dividir a informação complexa nas suas partes, analisar a relação entre as partes e o modo como formam um todo, assim como os princípios envolvidos ou erros subjacentes. Exemplos: Determinar os pontos principais de um artigo, história, regime político ou sistema económico; os postulados e consequências de uma teoria.
5. *Síntese*: Capacidade para formar e criar uma estrutura nova ou um todo a partir dos elementos constitutivos. Exemplos: Integrar informação de áreas diferentes para elaborar um plano, projecto, experiência ou resolver um problema.
6. *Avaliação*: Capacidade para avaliar e julgar a informação e os acontecimentos em relação a uma norma. Exemplos: Analisar os comportamentos de uma criança e avaliar se é ou não preciso uma intervenção especializada; avaliar a relação entre desenvolvimento e a preservação do ambiente; avaliar o melhor carro (ou casa, livro, etc.) de acordo com diferentes parâmetros.

Estes objectivos constituem para muitos uma hierarquia de objectivos educativos desenvolvendo-se desde os mais simples (conhecimento) até aos mais complexos (síntese e avaliação). Embora esta estrutura hierárquica não seja aceite por todos e haja dificuldades em aplicá-la em certas áreas como a matemática, é no entanto uma tipologia considerada aceitável para orientar e planear a elaboração de questões tanto em exames com formato de

² Bloom estabeleceu uma taxinomia sobre os objectivos da educação, que dividiu em três domínios: Cognitivo, afectivo e psicomotor. Estes objectivos estão ligados entre si, na medida em que um aluno ao tirar notas de uma aula, compreende o que ouve (cognitivo), sente agrado ou simpatia pelo conhecimento (afectivo) e escreve o que considera mais importante (psicomotor).

PEMs como de PDs. Em geral, a primeira e a segunda categorias de Bloom são consideradas como as mais fáceis de aplicar na elaboração de PEMs; veja-se alguns exemplos na II Parte.

Segundo Gronlund (1988) os primeiros quatro objectivos de Bloom podem ser realizados em exames com o formato de PEMs, verdadeiro-falso e emparelhamento, enquanto que os objectivos de síntese e avaliação seriam melhor avaliados em exames por PDs. As PDs seriam também indicadas para os restantes objectivos, nomeadamente os objectivos intermédios de aplicação e análise.

Memória versus raciocínio (#16)

A questão sobre a validação de PEMs que pretendem medir os processos de pensar superiores esbarra com algumas dúvidas, uma das quais é saber se os alunos seleccionam a resposta correcta com base num *raciocínio* complexo, ou antes com base na *memória* precisa de um conteúdo escolar anteriormente bem aprendido e estruturado. Por outras palavras, a resposta é seleccionada com base na familiaridade da resposta ou com base no resultado de raciocínios efectuados?

Esta pergunta é pertinente na medida em que os *peritos* na análise de questões complexas (por ex., especialistas médicos, peritos de acidentes, mestres de xadrez, grandes músicos e artistas) seleccionam muitas vezes a resposta certa com base na sua experiência prévia (memória), a qual exprime um grau elevado de estruturação de conhecimento, enquanto que os *principiantes* apenas acertam na resposta correcta — quando acertam! — depois de uma série de raciocínios intercalares. Aparentemente significa que os *peritos* usam a memória e os *principiantes* usam o raciocínio! Realmente o que pretendo afirmar é que muitas vezes não se sabe quais foram os processos mentais envolvidos na resposta correcta que um estudante seleccionou. Para a mesma resposta correcta a uma PEM e para alguns estudantes podem estar envolvidos processos mentais complexos, mas para outros estudantes podem ser processos mentais bem simples, nomeadamente a recordação de uma informação bem aprendida e estruturada (vide II Parte, pergunta 2).

Sistema psicológico subjacente (#17)

A avaliação dos objectivos educativos por PEMs foi criticada por ser um procedimento tipicamente behaviorista, cujo objectivo mais importante em termos educativos era o comportamento observável expresso pelo desempenho numa prova. Em contraste, o que deveria ser ressaltado em termos educativos eram as formas de expressão dos processos e operações de pensamento. Assim a perspectiva de avaliação deveria ser cognitiva. A perspectiva behaviorista aplicaria nas questões de avaliação termos como *definir*, *enumerar* e *calcular*, enquanto que a perspectiva cognitiva optaria por termos como *compreender*, *aplicar*, *comentar* e *interpretar*.

É importante ter em conta, no entanto, que o uso dos conceitos *definição* e *enumeração*, ou dos conceitos *compreensão* e *aplicação*, são nominalmente tão cognitivos como são behavioristas. Todos são behavioristas se se fixarem apenas no desempenho ou resposta independentemente do raciocínio subjacente. Todos são cognitivos se for possível inferir através do desempenho do estudante o maior ou menor grau de complexidade mental presente na forma como a resposta foi elaborada. É certo que esta inferência é mais fácil de se efectuar em exames de PDs do que em exames por PEMs, mas também é possível fazê-lo em PEMs quando o planeamento dos distractores a serem incluídos nas PEMs incluem conceitos e expressões erróneas frequentemente dadas em respostas passadas.

Orientações na Elaboração de PEMs

O processo de elaboração de PEMs é longo, laborioso, requer conhecimentos, experiência, bastante imaginação, revisões frequentes e o parecer e ajuda de colegas da mesma área de especialidade. Uma PEM, considerada por um autor como muito boa e imaginativa, revela-se deficiente, quando submetida à apreciação de um colega, devido a ambiguidades e presença implícita de pistas orientadoras. Em 1994 passei à volta de 5 semanas para elaborar e redigir cerca de 200 PEMs a partir de um manual escolar que tinha acabado de completar na área da aprendizagem e memória. Nos cinco anos lectivos seguintes cerca de 15% destas PEMs foram revistas ou anuladas devido a dúvidas dos estudantes ou reavaliação pessoal.

O formato mais convencional das PEMs é constituído por um *enunciado*, normalmente em forma de pergunta, seguido por quatro alternativas, assinaladas pelas letras A, B, C, e D, uma em cada linha numa distribuição vertical. Uma das alternativas é *correcta*, que deve ser identificada e assinalada pelo aluno, e as restantes alternativas são falsas e designam-se por *distractores*, na medida em que distraem o estudante, afastando-o da resposta correcta.

Há certas regras frequentemente seguidas para a elaboração de PEMs. No entanto estas regras não são o resultado de uma ciência própria, que não há. O que há são antes regras que têm a ver com a experiência, conselhos e sugestões de autores de manuais escolares e ainda com alguns princípios seguidos na elaboração de testes padronizados de aptidão escolar. Estas regras e sugestões são apresentadas em seguida e relacionam-se com a elaboração do enunciado, a resposta correcta, a escolha dos distractores e o formato global da pergunta.

O enunciado

O enunciado deve incluir a ideia central e a maior parte da informação em análise. O enunciado deve ser uma pergunta positiva. Os enunciados positivos são mais fáceis de compreender do que os negativos. No caso de um enunciado negativo recomenda-se que a palavra não seja sublinhada. A avaliação não tem por objectivo confundir os alunos ou passar-lhes rasteiras, antes pretende avaliar os conhecimentos adquiridos e os raciocínios efectuados.

O enunciado deve seguir de preferência um formato de pergunta em vez de um formato de completção. O enunciado deve formular a pergunta de tal modo que o estudante saiba responder, ou o possa fazer, antes de ler as alternativas. O formato em pergunta permite ao estudante evocar a resposta, imediatamente antes de ler as alternativas e assim evitar parte da confusão que a leitura destas pode gerar.

O formato de completção envolve maior sobrecarga de memória a curto prazo, obriga a leituras sucessivas, aumenta a dificuldade de compreensão por parte dos estudantes de culturas e etnias não-dominantes e é susceptível de originar maior ansiedade. O uso de espaços em branco nas perguntas de completção deve ser incluído de preferência no final do enunciado e não no início ou no meio. Um espaço em branco para completar no início do enunciado dificulta mais a compreensão.

A escolha da resposta correcta

No formato convencional de PEMs deve haver uma e só uma alternativa correcta. Às vezes por desconhecimento, distração, pressão do tempo, ambiguidade e falta de revisão há mais do que uma opção correcta (vide II Parte, perguntas 5, 6, 7, 9 e 10). Outras vezes uma alternativa, considerada correcta num momento, pode vir a tornar-se errada numa altura posterior, ou vice-versa, devido aos avanços do saber.³

Quando há fortes dúvidas de que uma alternativa é realmente correcta, o melhor a fazer é anular a questão. Às vezes pode ser aceitável reformular o enunciado perguntando qual é a *melhor* resposta ou justificação de entre as alternativas seguintes.

A resposta correcta deve distribuir-se de forma mais ou menos equiprovável pelas várias opções ou alternativas adoptadas no conjunto das PEMs que constituírem o exame. Se não for equiprovável, um estudante habilidoso é muito capaz de detectar o sistema de respostas.

Escolha dos distractores

Na elaboração de PEMs, o número de distractores, alternativas ou opções varia normalmente de três a cinco, sendo o número mais comum de quatro distractores. Na avaliação por PEMs, uma das maiores dificuldades é a selecção das alternativas ou distractores plausíveis e adequados. É difícil escolher bons distractores, mesmo por parte de autores de manuais escolares com longa experiência neste tipo de formato avaliativo.

Muitas PEMs permitem a elaboração de um bom distractor, às vezes dois distractores e mais raramente três. Quando os distractores não têm o mesmo poder de atracção em concorrência com a opção correcta, os distractores perdem eficácia e são facilmente rejeitados por um estudante minimamente competente na realização de testes de PEMs. Como consequência aumenta a probabilidade de acertar. Com distractores deficientes, a probabilidade de acertar na resposta correcta sobe progressivamente de 0.25 para 0.5 ou até mesmo para 1. por pergunta.

Como regra geral, um distractor deve ser seleccionado por estudantes com poucos ou nenhuns conhecimentos e rejeitado pelos bons alunos. Um distractor deve assemelhar-se a uma resposta correcta e gerar dúvidas relevantes naqueles estudantes que não possuem um conhecimento específico. Distractores adequados e eficazes têm origem muitas vezes em respostas erradas previamente dadas a perguntas abertas.

Na selecção dos distractores, os peritos desaconselham a inclusão da alternativa “Todas as anteriores [alternativas] são correctas”, ou “nenhuma das anteriores [alternativas] são correctas”, ou “não sei”. No caso de “Todas são correctas” em PEMs-4, basta que o estudante detecte que as duas primeiras alternativas são correctas para imediatamente considerar que a terceira alternativa também pode sê-lo, orientando-o para a resposta certa “Todas são correctas”. A alternativa “nenhuma é correcta” deve ser também evitada, porque para cada pergunta de exame deve haver sempre uma resposta correcta, se for convenientemente formulada. Deve-se evitar ainda as alternativas implausíveis, ou que incluem expressões absolutas como “sempre” e “nunca”, ou metade das alternativas a

³ É por esta razão que eu alerto os professores universitários, que estão mais próximos do estado actual do saber no âmbito da sua especialidade, a terem cautela com as explicações que informalmente dão a alunos do ensino secundário, onde os conteúdos escolares descritos nos manuais não têm o mesmo grau de actualização em relação ao que deve verificar-se na universidade.

começar por “É falso que ...” e a outra metade a começar por “É verdade que ...”, aumentando assim a probabilidade de acertar (Vide II Parte, pergunta 7).

É ainda aconselhável suprimir todas as pistas ou indicações de natureza gramatical, sintáctica e linguística que possam orientar o estudante para a alternativa correcta. Uma pista sintáctica relaciona-se com a extensão das frases de cada distractor. Esta extensão deve ser constante na medida do possível, ou na impossibilidade de o ser, dispor a extensão dos distractores do mais breve para o mais longo. Um pista linguística observa-se por exemplo na situação em que três distractores são termos típicos de uma língua, por exemplo a francesa, e o quarto distractor é típico de outra língua, por exemplo a inglesa. Se o enunciado se refere a um termo que tem uma configuração próxima da ortografia inglesa, não será difícil suspeitar qual será a escolha preferida. Quando todos os distractores incluem alguns conceitos comuns e outros diferentes, a ordem dos conceitos comuns em cada alternativa deve ser a mesma para evitar a sobrecarga de memória no processo de raciocínio (Vide II Parte, perguntas 5 e 8). Estas propostas e recomendações de avaliação têm por objectivo privilegiar os conhecimentos escolares adquiridos em detrimento da competência específica para a realização de provas deste tipo.

Formato global das PEM

Cada PEM deve estar relacionada com uma aprendizagem específica ou objectivo de instrução, sublinhando o alinhamento entre ensino e avaliação. Deve ainda basear-se em conhecimentos sólidos e bem fundamentados, raramente em perspectivas controversas. Se a perspectiva for controversa, o enunciado da pergunta deve assinalá-lo, tipo: “Segundo a perspectiva de ...” Cada PEM deve basear-se de preferência numa única ideia, ou tipo de conteúdo, seja conceito, facto, princípio ou procedimento.

As PEMs devem ser tanto quanto possível breves e curtas de forma a melhorar a clareza de interpretação e aumentar o número de perguntas do exame por período de tempo, obtendo-se deste modo um grau maior de fidelidade da prova. O vocabulário deve estar ao nível da amostra dos alunos que se pretende avaliar. Se for mais complexo do que a idade da amostra, resta saber se o resultado final representa a avaliação de conhecimentos ou a compreensão de leitura.

Na elaboração de PEMs são ainda de evitar as situações seguintes: questões problemáticas e ambíguas que mesmo os melhores alunos têm dificuldades em responder; a avaliação de conhecimentos irrelevantes e sem importância e informações que não foram aprendidas; discriminações bastante subtis; questões com a intenção de confundir e enganar os alunos.

As PEMs devem avaliar formas cada vez mais elevadas e complexas de raciocínio e pensamento em vez de formas mais simples, como a memória de factos, que constituem as questões mais fáceis de elaborar. O que é relevante na aprendizagem de um tópico, área, capítulo ou parte da matéria escolar deve orientar a elaboração das PEMs. É preciso ter em conta também o grau de especificidade de conhecimentos que se pretende avaliar. A ultra-especificidade das PEM pode violar a representatividade dos conhecimentos da disciplina escolar, ao inquirir por exemplo sobre a informação presente numa nota de rodapé, ou sobre o título completo de um artigo científico, ano e volume onde apareceu.

É aconselhável proceder-se a uma revisão periódica das PEMs em termos de clareza e de objectivos. Mesmo os peritos na redacção deste tipo de questões têm de abandonar com o decorrer do tempo um número significativo de PEMs por se revelarem confusas, ambíguas ou apresentarem falta de correspondência entre ensino e avaliação. Finalmente há provas e

testes de PEMs sujeitos a “copyright”, cuja cópia não autorizada representa uma infracção e constitui plágio. Para informações complementares sobre a elaboração de provas por PEM veja-se Haladyna (1994); Gronlund (1988); Airasian (1994) e McKeachie (1994).

Conclusão

A avaliação escolar por PEMs tem algumas limitações, mas se estas forem consideradas na elaboração de um exame, tendo em conta os objectivos escolares que se pretende promover, é um método bastante válido. A avaliação escolar exclusivamente por PEMs é pouco frequente no sistema escolar português, porque o processo de elaboração do exame é complexo e difícil, as questões passam para o domínio público e esgotam-se rapidamente. Para obviar a estes inconvenientes, que são marginais em relação à validade interna e externa deste formato, alguns professores usam uma mistura de PEMs e de PDs no exame, como aconteceu com os exames nacionais de psicologia do ensino secundário no ano 2000, de que a II Parte deste artigo reproduz algumas questões. Quer se considere isoladamente quer em conjunto com PDs, é de sublinhar que a avaliação exclusivamente por PEMs é um processo de avaliação válido e a diferenciação dos alunos estabelecida por este processo não é substancialmente diferente da estabelecida pelo formato de PDs (e.g., Bennett et al., 1990).

Penso que uma das diferenças principais entre uma avaliação por PEMs e por PDs relaciona-se com a expressão da criatividade dos estudantes. No entanto a criatividade revela-se sob muitas formas, e a forma mais fácil de ser expressa e avaliada é ao longo do ano lectivo em trabalhos e intervenções escolares mais do que no momento do exame final. Não me parece ainda que seja imprescindível que todas as disciplinas escolares tenham por objectivo a medição da criatividade — que um exame por PDs revelaria mais facilmente — até porque o que é ou não criativo não é uma questão fácil de ser avaliada de forma imediata e objectiva.

Quadro 1: Factores comparativos entre provas objectivas constituídas por perguntas de escolha múltipla (PEMs) e provas menos objectivas formadas por perguntas de desenvolvimento (PDs).

	Comparação	PEMs (perguntas de escolha múltipla)	PDs (perguntas de desenvolvimento)
1	Número de questões	Elevado	Reduzido
2	Cobertura dos tópicos escolares	Cobertura extensa	Cobertura reduzida, mas profunda
3	Tempo de elaboração das questões	Longo e difícil	Rápido e fácil
4	Custos financeiros	Elevados	Baixos
5	Fraude e cópia alheia	Fácil	Difícil
6	Tempo para realizar a prova; Objecções dos alunos	Suficiente e previsível; Sem objecções	Insuficiente e mais difícil de cumprir; Objecções frequentes.
7	Requer manual escolar único?	É aconselhável nas ciências sociais e humanas, excepto na categoria “conhecimento” de Bloom	Não, mesmo nas ciências sociais e humanas
8	Facilidade de correcção	Fácil e rápida	Difícil e longa
9	Fidelidade de correcção	Fiel e objectiva	Menos fiel; maior subjectividade
10	Causa alternativa dos erros dos alunos na realização da prova	Capacidade de leitura; Adivinhação	Capacidade de escrita; Maior facilidade de “enganar” o docente, escrevendo futilidades e “atirando o barro à parede”
11	Causa dos erros dos docentes na correcção da prova; Efeito	Problemas de atenção; Efeito insignificante	Dificuldade de concentração e sobrecarga de memória, atracção ilusória por uma escrita bonita e arranjo visual das respostas; Efeito significativo
12	Variância ou dispersão dos resultados	Maior em torno da média; Valores mais extremos	Menor em torno da média; Valores menos extremos
13	Cotação de cada resposta	Total e independente das precedentes	Parcial e dependente das precedentes
14	Prova de memória	Reconhecimento	Evocação
15	Medição das categorias de Bloom sobre objectivos cognitivos	Conhecimento, compreensão, aplicação e análise (1-4)	Todas (1-6), nomeadamente análise, síntese e avaliação;
16	Efeitos prováveis na aprendizagem dos alunos	Encoraja a memória, a identificação e a análise dos materiais escolares	Encoraja os estudantes a organizar, integrar e a exprimir os materiais de forma personalizada
17	Sistema psicológico subjacente	comportamentalista-beaviorista (?); baseado na memória e resposta dada	cognitivo (?); baseado no raciocínio e processos mentais superiores
18	Medição da criatividade	Inadequado	Adequado
19	Género favorecido	Equivalência de géneros	Feminino, devido à maior fluência verbal, compreensão e escrita; escrita mais atractiva (?)
20	Heterogeneidade cultural: avaliação	Mais favorável a meios heterogéneos	Menos favorável, porque implica o domínio correcto da língua e da escrita
21	Preferência dos alunos	Preferido, se não houver penalização; Ansiedade prévia menor; A ignorância tem custos menores	Preferência menor; Ansiedade prévia maior; A ignorância tem custos maiores
22	Deteção de erros perigosos e protecção pública	Inclui distractores perigosos que podem ser erradamente escolhidos; Maior protecção	Não detecta erros perigosos de forma sistemática; Menor protecção
23	Divulgação pública; Esgotamento da base de questões	Divulgação interdita ou limitada; Esgotamento rápido	Divulgação aceitável; Esgotamento demorado e longo
24	Copyright ©; Plágio	Passível de copyright ©; Tentação de plagiar	Copyright © muito raro; Fuga ao plágio através de paráfrases

Variáveis: (1 - 7) Organização; (8 - 13) Correcção; (14 - 21) Cognição e Preferências; (22 - 24) Outros;

II Parte

Esta II Parte analisa e comenta 10 tipos de perguntas de escolha múltipla (PEMs, 1-10), mais uma pergunta de desenvolvimento (PD, 11). As duas primeiras PEMs referem-se a uma questão específica respeitante ao formato das PEMs (1) e a outra questão sobre a prevalência ou não da memória versus raciocínio na resposta a uma PEM (2). As restantes oito PEMs (3-10) e a pergunta de desenvolvimento (PD 11) são perguntas que foram incluídas nos exames nacionais de psicologia do ensino secundário no ano 2000. Todas estas questões são seguidas de um comentário crítico, que analisa aspectos formais e aspectos de conteúdo. O objectivo deste comentário é mostrar a dificuldade de elaboração de enunciados precisos e distractores de qualidade na avaliação escolar por meio de PEMs e a necessidade de uma revisão atenta e periódica a ser efectuada por especialistas. Gostaria de sublinhar que não me move qualquer animosidade ou falta de apreço pelo júri que elaborou estas questões, que aliás desconheço. Se pretendesse criar polémica em vez de iniciar um debate científico-pedagógico sobre este tema, teria enviado na altura dos exames nacionais uma carta para os jornais diários, tão ávidos deste tipo de disputas.

1 Questão PEM, exemplo de um tipo de emparelhamento

Relacione os investigadores referidos na coluna da esquerda com as áreas psicológicas da direita:

A.	Piaget	1.	Psicologia social
B.	Skinner	2.	Psicologia do desenvolvimento
C.	Kurt Lewin	3.	Psicologia cognitiva
D.	Neisser	4.	Psicologia da aprendizagem

Resposta correcta: A2-B4-C1-D3. Esta questão é um exemplo de um tipo de PEM, em que as alternativas são respostas correctas que devem ser emparelhadas pelo estudante. Vantagens: Correção objectiva e obtenção de uma quantidade superior de informação em relação a uma PEM convencional. Enquadra-se no objectivo 2 de Bloom.

2 Memória vs raciocínio; Peritos vs principiantes

Qual é a maior via rodoviária da cidade do Porto?

- A. Boavista
- B. Marginal
- C. Via de cintura interna
- D.* Circunvalação

Nota: Para responder a esta questão, um principiante precisa de ter conhecimentos e saber princípios de geometria. Mas um perito, que pode ser qualquer habitante do Porto com alguma experiência de condução, apenas precisa de fazer apelo à memória. O principiante tem de ter o conhecimento de memória ou suspeitar que a circunvalação é a cintura exterior que rodeia a cidade do Porto e define os seus limites por terra. E deve ter ainda armazenado na memória princípios de geometria (comparação entre circunvalação e VCI) baseados no conhecimento de que quanto maior for o raio de um círculo maior será o seu perímetro. Uma pessoa bastante familiarizada com estas vias (perito), selecciona imediatamente a resposta correcta a partir de memória. Apenas os

principiantes têm necessidade de efectuar toda a sequência de raciocínios envolvidos. Esta distinção entre peritos e principiantes ajuda a compreender melhor as diferenças de pensar entre aqueles que têm um conhecimento organizado e estruturado e aqueles outros que não têm um conhecimento memorizado e precisam de usar informações dispersas e incertas para seleccionar a resposta correcta.

3⁴*ES/psic 140-2000/Modelo/PEM-5.1*

Segundo **Piaget**, os estilos de **pensamento intuitivo e mágico** e o **jogo simbólico** são característicos do estágio:

- A. das operações formais;
- B. das operações concretas;
- C.* pré-operatório;
- D. sensório-motor.

Nota: PEM formalmente correcta. Enquadra-se no objectivo 1 de Bloom.

4

ES/psic 140-2000/Modelo/PEM-4.1

Os **processos de liderança**:

- A. são característicos de qualquer grupo; são processos de conformismo;
- B. são característicos dos grupos informais; são processos de cooperação;
- C. são característicos dos grupos formais; são processos de obediência;
- D.* são característicos de qualquer grupo; são processos de influência social.

Nota: PEM formalmente incorrecta, porque a expressão “são característicos de” deveria fazer parte do enunciado. Provoca sobrecarga de memória e leituras desnecessárias. Objectivo 2 de Bloom.

5

ES/psic 140-2000/Modelo/PEM-7.1

O **processo de aprendizagem** é influenciado por:

- A.* inteligência; motivação; aprendizagens anteriores;
- B. motivação; factores sociais; anamnese;
- C. inteligência; motivação; orientação sexual;
- D. motivação; aprendizagens anteriores; anamnese.

Nota: PEM formalmente correcta, mas incorrecta em termos de conteúdo. Não há uma resposta correcta exclusiva, mesmo a opção C. pode ser considerada correcta. Em Portugal no ano de 1995/96 havia 130 raparigas por 100 rapazes no ensino superior (Eurostat) e nesse ano por cada 100 licenciados, 37 eram rapazes e 63 raparigas. Este padrão verifica-se à distância noutros países. Quem pode afirmar que o sexo não conta? A pergunta sugere ainda que há apenas um processo de aprendizagem. Mas não há um, há vários (e.g., Pinto, 2001b). A aprendizagem da ameba, da aplysia e da mosca do vinagre é influenciada porquê? Pela inteligência? Irrelevante; seria esvaziar o conceito de

⁴ A expressão ES/psic 140-2000/1F.1Ch.1v/PEM-1.1 significa Exame Nacional do Ensino Secundário / exame de psicologia nº 140 do ano 2000 / 1ª Fase. 1ª Chamada. 1ª versão / Pergunta de Escolha Múltipla - nº 1.1 do exame. Modelo refere-se a um Modelo de exame do ano 2000, previamente conhecido dos alunos.

inteligência. Esta PEM seria mais aceitável se o enunciado inclísse a palavra “principalmente” e “humano” assim: «O **processo de aprendizagem humano** é influenciado principalmente por:» Objectivo 2 de Bloom.

6

ES/psic 140-2000/Modelo/PEM-10.1

A **personalidade** é uma **construção pessoal** que, ao longo da vida:

- A. é estável, uma vez que resulta da interacção dinâmica entre agentes de socialização durante a infância;
- B. é passível de mudanças, uma vez que resulta da maturação e do desenvolvimento do sistema nervoso;
- C.* é passível de mudanças, uma vez que resulta da interacção dinâmica entre o indivíduo e o meio;
- D. é estável, uma vez que resulta dos traços de personalidade herdados dos progenitores.

Nota: PEM formalmente correcta, mas com uma redução discutível das opções apenas a duas (estável vs mudança). Não há uma resposta correcta exclusiva. A resposta correcta depende do modelo teórico subjacente de personalidade: As teorias dos traços defendem a estabilidade; as teorias behavioristas e situacionistas defendem a mudança; as teorias interaccionistas defendem a estabilidade ou mudança conforme as situações. O enunciado da PEM supõe que há uma teoria de personalidade melhor do que todas as outras em termos explicativos e com apoio empírico sólido, mas actualmente não há! A PEM seria mais aceitável se o enunciado fosse: «Segundo a teoria interaccionista, a personalidade é:» e usando distractores diferentes. Objectivo 2 de Bloom.

7

ES/psic 140-2000/1F.1Ch.1v/PEM-10.2

Os factores determinantes na construção da personalidade são os factores hereditários. Esta afirmação é:

- A. falsa, porque os factores hereditários interagem com os factores orgânicos e com as experiências familiares;
- B. verdadeiro, porque a hereditariedade determina a forma como o organismo interage com o meio;
- C.* falsa, porque os factores hereditários interagem com os factores ambientais e com as experiências pessoais;
- D. verdadeiro, porque a hereditariedade determina os factores orgânicos de relação com o meio.

Nota: O júri considerou correcta a opção C; Mas segundo os estudos dos geneticistas comportamentais realizados nos últimos 10 anos com gémeos monozigóticos (e.g., Bouchard et al., 1990; Bouchard, 1994) e de acordo com o modelo de desenvolvimento de Judith Harris (1995) a opção correcta estaria mais próxima de B. Comentário: Um saber actualizado pode ser contraproducente. Veja-se a nota de rodapé 3.

8

EES/psic 140-2000/1F.1Ch.1v/PEM-5.1

Segundo **Piaget**, o **desenvolvimento cognitivo** é influenciado por **factores** como:

- A.* hereditariedade; experiência física; transmissão social; equilibração;
- B. maturação; acomodação; estádios de desenvolvimento; seriação;
- C. acção sobre os objectos; equilibração; egocentrismo; esquemas;
- D. adaptação; estádios de desenvolvimento; coordenação; classificação

*Nota: PEM formalmente correcta, mas de reflexão longa devido à sobrecarga de memória envolvida, nas várias comparações entre termos e expressões referidos nas 4 alternativas. Além de conhecimentos sobre os factores que influenciam o **desenvolvimento cognitivo** para Piaget, esta pergunta faz apelo à capacidade da memória a curto prazo e a habilidades específicas para realizar testes de PEMs. Seria preferível manter um primeiro termo comum às 4 opções, mais um segundo termo comum à 1ª e 2ª opções e um terceiro termo comum à 3ª e 4ª opções, mais 4 termos diferentes em último lugar em cada opção. Este problema observa-se ainda na questão ES/psic 140-2000/1F.1Ch.1v/PEM-7.1*

9

ES/psic 140-2000/1F.1Ch.1v/PEM-2.1

O **método experimental** apresenta, entre outras, dificuldades em:

- A.* controlar as expectativas dos participantes; isolar a variável independente;
- B. criar situações experimentais adequadas; manipular as variáveis internas;
- C. neutralizar os efeitos do experimentador; manipular as variáveis externas;
- D. aplicar procedimentos técnicos variados; isolar a variável dependente.

Nota: Não há uma resposta correcta convincente. Com excepção da opção D, as opções A, B e C são respostas correctas possíveis. Além disto as dificuldades A, B e C não são específicas do método experimental, podendo aplicar-se também ao método diferencial, e até a outros (Pinto, 1990). Este problema observa-se ainda na questão seguinte nº 10.

10

ES/psic 140-2000/1F.1Ch.1v/PEM-9.1

Globalmente, a **inteligência** pode caracterizar-se como a capacidade de:

- A. utilizar a linguagem fluentemente;
- B.* interagir e adaptar-se ao meio;
- C. armazenar conhecimentos;
- D. resolver problemas abstractos.

Nota: Não há uma resposta correcta convincente. A opção B considerada correcta também caracteriza a percepção visual, a aprendizagem, a emoção, a linguagem, etc. O que é afinal único e característico da inteligência que a diferencia dos restantes processos cognitivos? A alternativa correcta B. é uma hipótese, talvez a melhor, mas é uma definição entre várias. Por exemplo, se se aceitar o modelo de inteligências múltiplas de Howard Gardner (1993), muito popular em certos meios, não há alternativa correcta, porque todas as quatro alternativas são correctas ao representarem diferentes inteligências.

4. A personalidade, ainda que se refira a padrões de comportamento relativamente estáveis, não é imutável. Explique o processo dinâmico de construção da personalidade.

Nota: Pergunta de desenvolvimento (PD). Objectivo ambíguo. O que se pretende com esta pergunta? (1) Descrever apenas os factores responsáveis pela formação da personalidade? (2) Ou explicar a formação da personalidade segundo o modelo dinâmico de Freud? Parece que a resposta esperada pela correcção proposta num dos jornais diários era mais do tipo (1). No entanto o termo “explicar” subentende o recurso a uma teoria ou modelo para explicar a personalidade, deixando subentender que se pretendia uma resposta que tivesse em conta o modelo dinâmico de Freud.

Referências

- Airasian, P. W. (1994). *Classroom assessment* (2ª ed.). New York: McGraw-Hill.
- Anderson, J. R. (1990). *The adaptative character of thought*. Hillsdale, NJ: LEA.
- Baddeley, A. D. (1997). *Human memory: Theory and practice (Revised edition)*. Hove, UK: Psychology Press.
- Bennett, R. E., Rock, D. A., e Wang, M. (1990). Equivalence of free-response and multiple-choice items. *Journal of Educational Measurement*, 28, 77-92.
- Bloom, B. S., Engelhart, M. B., Furst, E. J., Hill, W. H., e Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: The cognitive domain*. New York: Longmans Green.
- Bouchard, T. J. (1994). Genes, environment, and personality. *Science*, 264, 1700-1701.
- Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., e Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*, 250(4978), 223-228.
- Cohen, N. J., e Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, 210, 207-210.
- Gagné, R. M. (1968). Learning hierarchies. *Educational Psychologist*, 6, 1-9.
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. New York: Basic Books.
- Gronlund, N. E. (1988). *How to construct achievement tests* (4 ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Haladyna, T. M. (1994). *Developing and validating multiple-choice test items*. Hove, East Sussex: LEA.
- Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review*, 102, 458-489.
- McKeachie. (1994). *Teaching tips: Strategies, research, and theory for college and university teachers* (9ª ed.). Lexington, MA: D. C. Heath.
- Miller, H., Williams, R. G., e Haladyna, T. M. (1978). *Beyond facts: Objective ways to measure thinking*. Englewood Cliffs, NJ: Educational Technology.
- Pinto, A. C. (1990). *Metodologia da investigação psicológica*. Porto: Jornal de Psicologia.
- Pinto, A. C. (1992). *Temas de memória humana*. Porto: Fundação Eng. António de Almeida.
- Pinto, A. C. (1998). Long term retention of school contents on Portuguese literature and geography. *Studia Psychologica*, 40, 219-225.

- Pinto, A. C. (2001a). Memória, cognição e educação: Implicações mútuas. In B. D. C. Cunha (Ed.), *Educação, cognição e desenvolvimento*. No prelo.
- Pinto, A. C. (2001b). *Psicologia geral*. Lisboa: Universidade Aberta (no prelo)
- Royer, J. M., Cisero, C. A., e Carlo, M. S. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research*, 63, 201-243.

MULTIPLE-CHOICE TESTS: SOME FORMAT AND VALIDITY ISSUES

Amâncio da Costa Pinto

Faculty of Psychology and Educational Sciences, University of Porto, Portugal

Abstract: Multiple-choice tests are becoming more and more popular in Portugal in final examinations. They are more objective and easy to score than essay tests. However these test formats have strengths and weakness. The aim of this paper is to describe and examine some major factors involved in developing multiple-choice test itens, namely issues related to format and writing, errors and validity, student preferences and cognitive categories implied. To illustrate most of these issues, Part II of this paper includes a list of several multiple-choice test items, most of them were selected from the Psychology National Exams given to students in the final year of the secondary education in Portugal in 2000.

Key-Words: *Multiple-choice test, assessment, psychology, national exams, secondary education.*