

MINISTÉRIO DA EDUCAÇÃO
DEPARTAMENTO DO ENSINO SECUNDÁRIO

Avaliação e Desempenho

Texto de Apoio

Setembro 2001

AVALIAÇÃO DE DESEMPENHO

▪ Da avaliação tradicional à avaliação de desempenho

A avaliação tradicional centra-se na verificação das aprendizagens realizadas pelos alunos, predominantemente no domínio cognitivo. A avaliação tem como principal função “medir” os conhecimentos adquiridos pelos estudantes, a fim de lhes atribuir uma classificação. Acredita-se que é possível avaliar, de forma objectiva, os resultados das aprendizagens dos alunos, utilizando instrumentos rigorosos de medida: os testes objectivos.

Contudo, estudos realizados por diversos investigadores - Broadfoot (1994), Fernandes (1994) e Quellmalz (1985), entre outros – apontaram como uma limitação significativa dos testes a ênfase na avaliação de aspectos mais facilmente mensuráveis (conhecimento e compreensão), em detrimento de processos complexos de pensamento. Ora, numa sociedade complexa como a nossa, os estudantes têm de desenvolver processos que lhes permitam atingir níveis complexos de pensamento - processos cognitivos de análise, comparação, inferência e avaliação - relacionados com a resolução de problemas e o pensamento crítico (Quellmalz, 1985; Resnick e Resnick, 1992). Mas para desenvolver e avaliar tais processos cognitivos de nível superior é indispensável que os estudantes desempenhem tarefas de aprendizagem próximas das tarefas “reais” que virão a desempenhar futuramente, pois são essas que convocam o desempenho intelectual complexo.

A crescente evidência de que os testes tradicionais - objectivos, de referência normativa, de escolha-múltipla - não possibilitavam a avaliação de capacidades cognitivas complexas, nem do desempenho dos alunos, incentivando um ensino tendencialmente descontextualizado e baseado na memorização, contribuiu de modo significativo para o reconhecimento da necessidade de uma avaliação “autêntica” ou “avaliação do desempenho” (Gardner, 1992; Murphy, 1995).

Numa sociedade em que se pretende assegurar o sucesso escolar a todos, as práticas pedagógicas devem ser dinâmicas e proporcionar oportunidades de aprendizagem a todos os alunos. A avaliação deve identificar dificuldades e sugerir formas de ajudar os alunos: tem como principais funções a compreensão e o melhoramento da prática educativa; deve enfatizar o controlo e o progresso individuais (identificação dos pontos fortes e das necessidades dos alunos e conseqüente adaptação do ensino, por parte do professor) constituindo o *feedback* constante fornecido aos alunos, um dos seus factores-chave. A ênfase desta avaliação deverá colocar-se na melhoria das aprendizagens dos alunos. Trata-se de uma concepção alargada de avaliação em que já não basta aplicar testes e exames aos alunos; tem de se apreciar comportamentos, conhecimentos, capacidades, atitudes, hábitos, interesses, de forma a assegurar informação que permita o desenvolvimento de um conjunto alargado e integrado de capacidades e competências.

▪ Avaliação de Desempenho – o que é?

A avaliação autêntica surge, assim, da procura de soluções mais justas e adequadas às situações reais de ensino-aprendizagem que ocorrem na sala de aula. Ao realizar tarefas de desempenho, os alunos utilizam, em simultâneo, tal como no “mundo real”, um conjunto de estratégias de aquisição e aplicação de (a) conhecimentos, (b) capacidades e competências processuais (processos complexos de pensamento, p.ex, resolução de problemas), e (c) hábitos de trabalho. Tais tarefas constituem, simultaneamente, oportunidades para aprender e oportunidades para avaliar as competências dos estudantes.

Em síntese, a avaliação *autêntica, contextualizada* ou *de desempenho*, é perspectivada como uma abordagem em que as tarefas de avaliação estão intimamente relacionadas com o desempenho desejado e se desenrolam num contexto autêntico. Tendencialmente, as actividades de aprendizagem são coincidentes com as actividades de avaliação: o aluno, ao realizar tarefas de desempenho - tarefas reais em que constrói respostas, cria produtos ou faz apresentações - evidencia os seus conhecimentos e competências, fornecendo, simultaneamente, elementos para a sua avaliação.

Este paradigma deriva da necessidade e da exigência de diferenciar a avaliação, de acordo com as capacidades, necessidades e interesses individuais, de modo a possibilitar a equidade, isto é, a igualdade de oportunidades de aprendizagem a todos os alunos, independentemente da sua proveniência sócio-cultural. Nesta perspectiva, é imprescindível respeitar a variabilidade individual quanto ao ritmo e ao estilo de aprendizagem, ou, como afirma Gardner (1992), torna-se fundamental atender à diversidade (diferenças individuais, níveis de desenvolvimento, tipos de aptidões...).

Este tipo de avaliação, contextualizada, privilegia técnicas que recolhem a informação no decurso da aprendizagem. Tais técnicas, como afirma Madaus (1994), já não requerem a selecção de respostas entre itens alternativos, elaborados por outrem, mas exigem que os alunos construam as suas próprias respostas, desempenhem tarefas (tarefas “do mundo real”, nas palavras de Wiggins, 1989), ou produzam algo que será avaliado, segundo critérios relevantes. Valoriza-se o desenvolvimento de capacidades produtivas e reflexivas, promovidas em projectos a longo prazo, objectivos bem patentes na metáfora usada por Maeroff (1991) ao comparar a avaliação autêntica com uma gravação em *video* com a duração de um semestre e o teste com uma *foto instantânea*, tirada com uma *Polaroid*, para captar apenas *aquele momento* da aprendizagem.

Os testes não são de excluir, mas assumem, nesta abordagem, um novo papel: os resultados do teste deixam de constituir dados inquestionáveis e preponderantes e passam a dar informação passível de influenciar a aprendizagem, não só através da orientação do professor, como da própria auto-avaliação e *auto-instrução* dos alunos (Glaser 1990). A informação pertinente sobre as aprendizagens realizadas pelos estudantes é agora recolhida através de um conjunto diversificado de técnicas e instrumentos.

Não há, porém, competição entre os testes tradicionais e a avaliação de desempenho. O desafio está em encontrar o equilíbrio entre eles: enquanto que o teste responde à pergunta “O que é que sabes?” e implica competências como *memorizar, identificar, listar e relacionar*, a avaliação do desempenho responde à questão “Em que medida és capaz de utilizar aquilo que sabes?” e mobiliza competências como *classificar, comparar, analisar e avaliar*.

De acordo com Murphy (1995), é necessário propor tarefas interactivas e diversificadas, que os alunos considerem como necessárias e significativas para se atingirem os objectivos pretendidos: avaliar e aprofundar capacidades individuais. Darling-Hammond (1994) inclui diversas tarefas no âmbito da avaliação autêntica: apresentações, debates ou exposições orais; recolha de produções escritas, gravações em *vídeo* (dos desempenhos e de outras situações de aprendizagem); resolução de problemas; experiências ou resultados de pesquisas; observações e inventários do trabalho e do comportamento individuais, bem como de trabalhos em grupo.

▪ **Avaliação de desempenho: como se implementa?**

- Escolha das tarefas/actividades

As tarefas (incluídas ou não em projectos) deverão ser seleccionadas de preferência com a colaboração dos alunos, pois devem ser significativas e susceptíveis de provocar a adesão e o envolvimento efectivo dos estudantes. Além disso, têm de ser escolhidas actividades relevantes que impliquem a aplicação de conhecimentos, capacidades e competências, de forma contextualizada, que possibilitem aos alunos a aquisição das aprendizagens pretendidas.

- Definição de critérios

A definição clara dos critérios de avaliação promove o rigor na avaliação, uma vez que o professor sabe exactamente quais as capacidades/competências e ou conceitos que está a avaliar, e os alunos sabem o que é esperado do seu desempenho.

Para definir critérios de avaliação, Airasian (1991, p.244) sugere os passos seguintes:

1. Identificar o desempenho ou tarefa a avaliar e desempenhá-la, por si próprio, ou, em último caso, imaginar que a está a desempenhar.
2. Listar os aspectos importantes do desempenho ou produto final.
3. Tentar limitar o nº de critérios de desempenho, de modo a que possam ser efectivamente observados durante o desempenho pelo aluno.
4. Se possível, ter outros professores a pensarem nos comportamentos importantes incluídos na realização da tarefa.
5. Expressar os critérios de desempenho, em termos de comportamentos observáveis dos alunos, ou de características dos produtos a elaborar.

6. Organizar os critérios do desempenho pela ordem em que eles provavelmente serão observados.

A definição dos critérios poderá ser feita com a participação dos alunos. Caso não o seja, os alunos deverão ter conhecimento prévio dos critérios que serão observados no desenvolvimento de cada tarefa de desempenho.

- Criação de rubricas de desempenho

Uma rubrica é um sistema de classificação pelo qual o professor determina a que nível de proficiência um aluno é capaz de desempenhar uma tarefa ou apresentar/evidenciar conhecimento de um conteúdo/conceito. Trata-se de explicitar, através de uma descrição detalhada, os níveis de qualidade de um desempenho ou de um produto. Com rubricas podem definir-se os diferentes níveis de proficiência de cada critério. O desempenho em cada nível deve ser claramente definido e traduzir rigorosamente o critério correspondente (Airasian, 1991; Popham, 1995; Stiggins, 1994).

Tal como sugerido para a definição de critérios, os alunos podem participar na elaboração das rubricas ou, pelo menos, ter conhecimento prévio das mesmas.

- Instrumentos de avaliação

Existem diversos instrumentos que podem ser utilizados para registo dos resultados da avaliação do desempenho (Airasian, 1991; Stiggins, 1994):

- Lista de Verificação – fornece apenas a indicação sobre a presença ou ausência de certos elementos no desempenho avaliado.
- Narrativa/Incidente Crítico – permite relatar, de forma descritiva, como decorreu o desempenho e explicitar em que medida os alunos atingiram os objectivos.
- Escala de Classificação – permite indicar o nível em que os objectivos foram atingidos; geralmente, os professores utilizam uma escala numérica. Por exemplo, pode-se classificar cada critério numa escala de 1 a 5, em que 1 signifique “competência quase inexistente” e 5 signifique “competência extremamente bem executada”.

É, ainda, possível, utilizar um tipo de instrumento que combine elementos dos anteriores, como é o caso das Listas de Avaliação do Desempenho nas Tarefas, que, simultaneamente: (a) fornece aos alunos uma estrutura que lhes permite trabalhar de forma independente e os encoraja a ter atenção à qualidade do trabalho; e (b) permite ao professor fornecer, de forma eficiente, informação aos alunos sobre os pontos fortes e fracos do seu trabalho, isto é, sobre o que precisam de saber e são capazes de fazer (exemplo em anexo).

- Atribuição da classificação

A avaliação sumativa final - de período, ano ou ciclo – deve traduzir-se, como é sabido, por uma classificação quantitativa utilizando uma escala de 0 a 20 valores. Tal como antes referido para a organização e implementação da avaliação de desempenho, também a decisão classificativa deve resultar de uma reflexão de critérios e acerto de procedimentos ao nível das estruturas pedagógicas dos professores: conselho pedagógico, departamento curricular, área disciplinar, conselho de turma.

Toda a avaliação integra, simultaneamente, por um lado, um juízo sancionador relativamente ao que foi já realizado e, por outro, uma inferência relativamente ao que se prevê que o aluno possa vir a realizar, com base no comportamento demonstrado.

A preocupação máxima deverá orientar-se no sentido de explorar as potencialidades *formativas* da classificação. Esta, enquanto expressão simbólica do resultado da avaliação, nas palavras de Zabalza (1995), constitui, para além da sua eventual função certificativa, mais uma forma ao dispor do professor para fornecer *feedback* ao aluno, integrando nela não só a expressão quantificada do juízo sumativo mas, igualmente, indicadores da confiança nas capacidades académicas e pessoais do aluno: potencialidades intelectuais que necessitam de contextos para desenvolvimento, comportamentos firmados de perseverança, confiança, responsabilidade e outros. Por isso, a atribuição criteriosa de uma classificação deve, segundo Messick (1989a), contemplar as consequências previstas para o aluno como aprendiz e como pessoa, isto é, deve promover a progressão sustentada do percurso escolar e pessoal do aluno.

Sugere-se, pois, que os critérios para atribuir classificações finais sejam debatidos e acordados pelos professores – a nível de disciplina e de turma – de forma a expressarem uma visão coerente e equitativa do desempenho global do aluno. Para tal, é importante acertar a relevância das diferentes componentes a integrar na avaliação, assim como a utilização de percentagens ou cotações ponderadas dos diferentes dados de avaliação disponíveis. Na avaliação de desempenho, conforme foi já referido, o produto final produzido pelo aluno não tem um valor autónomo face ao processo que o originou. Por isso é importante a diversificação de técnicas e de instrumentos de avaliação. Esta diversidade de dados acerca do comportamento do aluno deve ser considerada e integrada a quando da decisão classificativa.

O que deve, portanto, ser assegurado é a adequação da lógica classificativa às necessidades da avaliação de desempenho. Não será, por certo, admissível que o esforço concertado realizado por professores e alunos no sentido de desenvolver uma avaliação participada, contextualizada e alinhada pelo currículo seja comprometido por concepções e práticas de classificação individualistas, casuísticas e reveladoras de um excesso sancionador e comparativista, mais preocupadas com as subtilezas do pormenor do que com a riqueza do essencial: a promoção do desenvolvimento global do aluno através das aprendizagens adequadas.

▪ **BIBLIOGRAFIA**

- Airasian, P.W. (1991). *Classroom assessment*. N.Y.: McGraw-Hill.
- Broadfoot, P. (1994). Les résultats de l'enseignement. In OCDE (Ed.), *Évaluer l'enseignement: De l'utilité des indicateurs internationaux* (pp. 260-286). Paris: OCDE.
- Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 64, 1, 5-30.
- Gardner, H. (1992). Assessment in context: The alternative to standardized testing. In B. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction*. (pp. 77-120). Boston: Kluwer.
- Gipps, C. (1995a). *Beyond testing: towards a theory of educational assessment*. London: The Falmer Press.
- Gipps, C. (1995b). What do we mean by equity in relation to assessment. *Assessment in Education: Principles, Policy & Practice*, 2, 3, 271-282.
- Glaser, R. (1990). Toward new models for assessment. *International Journal of Educational Research*, 14, 5, 475-483.
- Harnisch, D. e Mabry, L. (1993). Issues in the development and evaluation of alternative assessments. *Journal of Curriculum Studies*, 25, 2, 179-187.
- Hibbard, K. M. et al. (1996). *A teacher's guide to performance-based learning and assessment*. Alexandria, VA: ASCD.
- Madaus, G. (1994). A technological and historical consideration of equity issues associated with proposals to change the nation's testing policy. *Harvard Educational Review*, 64, 1, 76-95.
- Maeroff, G. (1991). Assessing alternative assessment. *Phi Delta Kappan*, 73, 4, 273-281.
- Murphy, P. (1995). Sources of inequity : understanding students' responses to assessment. *Assessment in Education: principles, policy & practice*, 2, 3, 249-270.
- Popham, W. J. (1995). *Classroom assessment: What teachers need to know*. Needham Heights, MA: Allyn and Bacon.
- Quellmalz, E. (1985). Needed: Better methods for testing higher order thinking skills. *Educational Leadership*, 43, 2, 29-35
- Resnick, I. & Resnick, D. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. Gifford & M. C. O'Connor (Eds.), *Changing assessments: alternative views of aptitude, achievement and instruction*. (pp.37-75). Boston: Kluwer.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. NY: MacMillan.
- Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, 46, 7, 41-47.

Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. Phi Delta Kappan, May, 703-713.

www.ed.gov/databases/ERIC_Digests/ed423312.html

www.tensigma.org/pbl.defined.html

www.ascd.org/readingroom/books/hibbard96book.html

ANEXO

Lista de Avaliação do Desempenho na Tarefa de Realização de um Gráfico			
	Pontuação por item	PONTUAÇÃO	
		Classificação Obtida:	
Intervenientes:		O aluno	O professor
1. Utilização de um tipo de gráfico apropriado (barras ou linha...)			
2. Utilização apropriada de origem e intervalos em cada eixo			
3. Existência de um título para o gráfico que estabelece claramente a relação entre os eixos			
4. Utilização apropriada de uma escala em cada eixo			
5. Identificação clara dos eixos.			
6. A variável independente está colocada no eixo (X) e a variável dependente está no eixo (Y)			
7. Os dados estão inseridos com rigor			
8. Há uma legenda apropriada			
9. As linhas ou barras utilizam bem o espaço do gráfico			
10. São usadas técnicas adequadas (côr, textura ou etiquetas) para facilitar a leitura do gráfico			
11. O gráfico ocupa um espaço adequado na folha de papel			
12. O gráfico está claro e bem apresentado			
13. O gráfico é fácil de interpretar			
TOTAL:			