



IAVE INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.

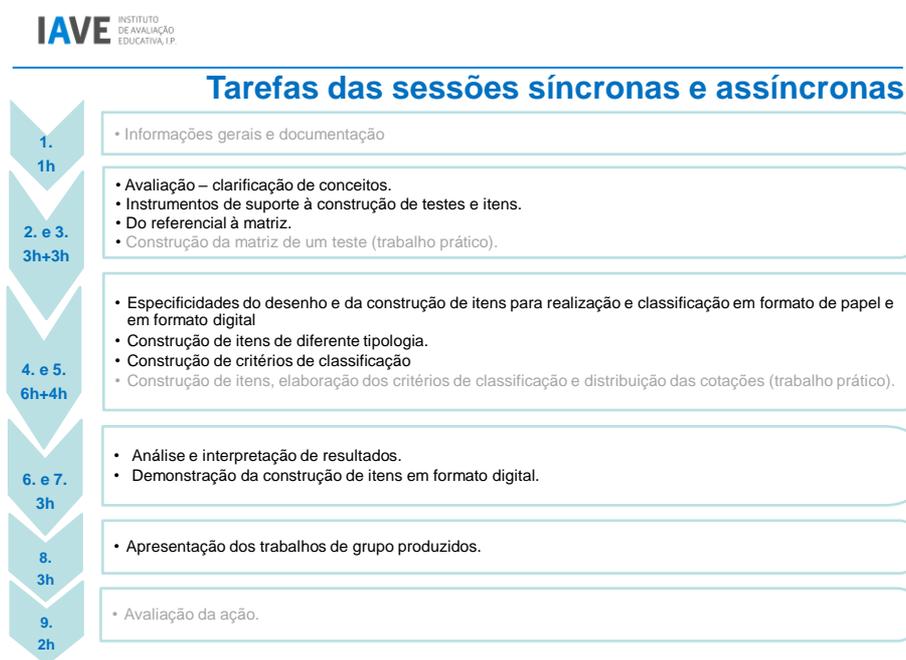
**Avaliar para aprender:
construção de instrumentos de avaliação**

Sessões 2 e 3

AprofGeo

Outubro/Novembro de 2022

1



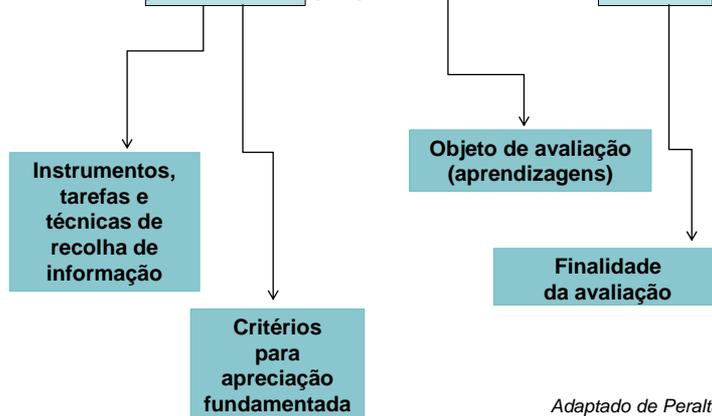
IAVE INSTITUTO DE AVALIAÇÃO EDUCATIVA, I.P.

Tarefas das sessões síncronas e assíncronas

1. 1h	• Informações gerais e documentação
2. e 3. 3h+3h	<ul style="list-style-type: none"> • Avaliação – clarificação de conceitos. • Instrumentos de suporte à construção de testes e itens. • Do referencial à matriz. • Construção da matriz de um teste (trabalho prático).
4. e 5. 6h+4h	<ul style="list-style-type: none"> • Especificidades do desenho e da construção de itens para realização e classificação em formato de papel e em formato digital • Construção de itens de diferente tipologia. • Construção de critérios de classificação • Construção de itens, elaboração dos critérios de classificação e distribuição das cotações (trabalho prático).
6. e 7. 3h	<ul style="list-style-type: none"> • Análise e interpretação de resultados. • Demonstração da construção de itens em formato digital.
8. 3h	• Apresentação dos trabalhos de grupo produzidos.
9. 2h	• Avaliação da ação.

2

A avaliação é a **recolha** sistemática de **informação** sobre a qual se formula um **juízo de valor** que permite a tomada de **decisões**.



Adaptado de Peralta, 2002

3

Que tipo de decisões (avaliar para quê)?

Com a finalidade de dar *feedback* para a melhoria

- ao aluno (regulação da aprendizagem)
- ao professor } (orientação do ensino)
- às escolas }

**Avaliação
PARA AS
aprendizagens**

Com a finalidade de fazer um *balanço das aprendizagens desenvolvidas* (informação sobre os resultados obtidos, certificação e seriação)

**Avaliação
DAS
aprendizagens**

The key difference between formative and summative assessment is not timing, but purpose and effect.

Gipps, 1994, p. 4

4

	Avaliação PARA AS aprendizagens	Avaliação DAS aprendizagens
Finalidade	Providenciar <i>feedback</i> para a melhoria. Orientar o ensino.	Proporcionar um juízo de valor sobre o objeto avaliado. Informar sobre a qualidade do objeto avaliado.
Uso	Guia para a tomada de decisão.	Determinar o sucesso ou insucesso; certificar.
Foco	Objetivos de aprendizagem; processos; estratégias.	Produtos; desempenhos.
Orientação	Pró-ativa.	Retroativa.
Momento	Em qualquer momento.	No final de um ciclo de aprendizagem.
Papel do erro	Formativo; informativo.	Punitivo; informativo.
Informação	Predominantemente qualitativa e descritiva.	Predominantemente quantitativa.

Adaptado de Stufflebeam & Shinkfield, 2007

5

Como recolher a informação?

Instrumentos, tarefas e técnicas de avaliação

- Concebidos e elaborados sob a responsabilidade da escola e dos professores.
- Concebidos e elaborados sob a responsabilidade da administração central.
 - Com impacto (*High stakes*) — Exames
 - Sem impacto (*Low stakes*) — Provas de Aferição

**Avaliação
interna**

**Avaliação
externa**

Todos podem proporcionar informação para *feedback* (mais ou menos descritiva)
É o propósito da avaliação que vai determinar a forma como são construídos

6

Como recolher a informação?

Instrumentos, tarefas e técnicas de avaliação

Inquérito	Observação	Análise	Testes
Entrevistas Questionários Técnicas sociométricas (...)	Registos de observação direta Listas de verificação (...)	Análise e sinalização de evidências Registos de incidentes críticos (...)	Testes escritos Testes de aptidão (...)
Instrumentos adequados ao tipo de informação que se pretende recolher			
Opiniões Perceções (...)	Desempenhos Produtos Interações Comportamentos (...)	Evidências relevantes durante o processo (...)	Objetivos de aprendizagem Desempenhos (...)

Adaptado de Ribeiro, 1991

7

Como formular juízos de valor?

Interpretando as respostas dos alunos por comparação a

- Norma de um dado grupo (avaliação **normativa**)
- Critérios pré-definidos (avaliação **criterial**)
- Desempenhos prévios do próprio indivíduo (avaliação **ipsativa**)

8

Como formular juízos de valor?

Critério

É um modo de interpretar a informação (Figari, 1996).

É uma forma de definir o sucesso (Wiggins, 1998).

É um modo de descrever as características (a qualidade) de um desempenho esperado (Brookhart, 2013).

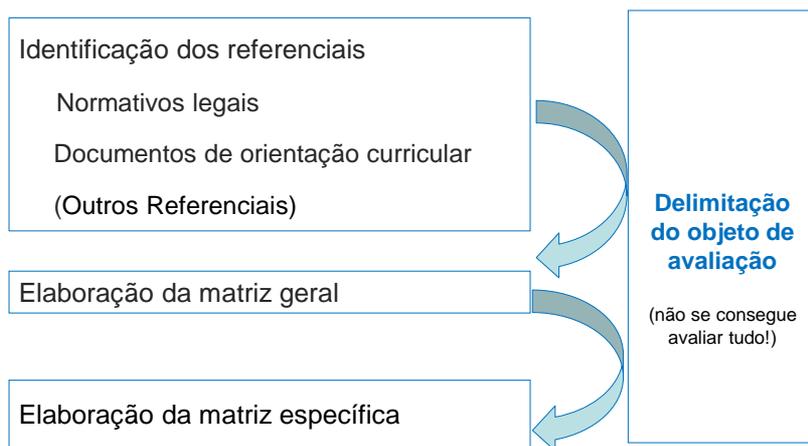
Definir critérios, para quê?

Para definir o que é o sucesso e perceber se foi atingido.

Para assegurar a qualidade da avaliação através de um processo transparente, participado e fundamentado.

9

Que informação recolher? Como definir o objeto em avaliação?



10

Percurso de um instrumento de avaliação externa: da conceção à tomada de decisões

Em função da finalidade da avaliação:

- Conceptualização e planeamento
 - Elaboração
 - Aplicação
 - Classificação
 - Análise de resultados
 - Elaboração de relatórios
 - Tomada de decisões
-

11

Conceitos centrais para a definição de um teste

- Um teste tem um **domínio** de interesse (corpo particular de conhecimentos, competências, capacidades, desempenhos).
- Um teste é uma **amostra** (de produtos, comportamentos, respostas, desempenhos).
- Um teste permite fazer **inferências** sobre o domínio e usá-las para descrever, tomar decisões ou determinar consequências relativamente aos avaliados.
- O grau em que essas inferências específicas são adequadas designa-se **validade**.

Madaus, Haney & Kreitzer, 2002

12

Qualidade da Avaliação

Validade

A informação recolhida com um dado instrumento é válida se essa informação for uma medida daquilo que se pretende medir.

Está relacionada com a adequação das interpretações dos resultados.

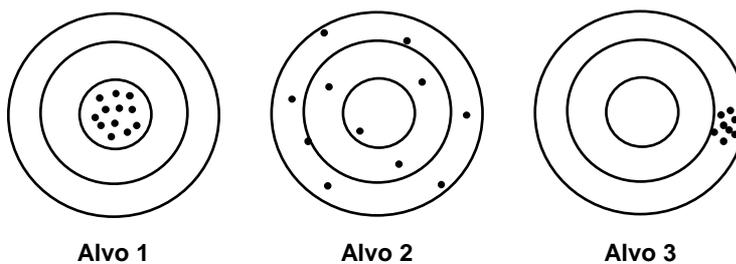
Fiabilidade

A fiabilidade da informação recolhida está associada à consistência dessa mesma informação.

Está relacionada com a uniformidade e a estabilidade dos resultados.

A falta de **fiabilidade** pode comprometer a **validade**.

13



Analogia entre a informação recolhida com um instrumento de avaliação e a posição num alvo dos tiros disparados por uma arma.

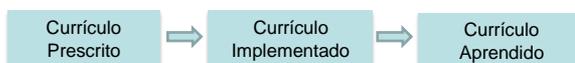
Miller, Linn & Gronlund, 2009

14

Fatores que afetam a validade e a fiabilidade

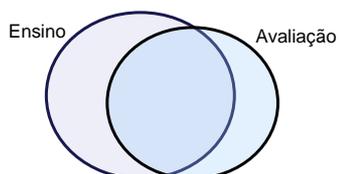
Na conceção e planeamento de um teste

- Definição incorreta do domínio



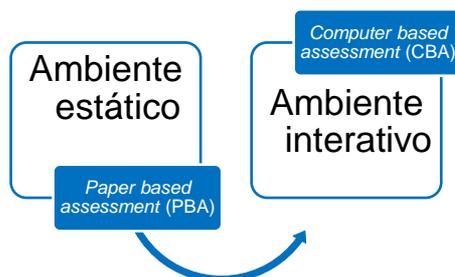
Outros modeladores do currículo: manual, professor (experiência pessoal e profissional)...

- Falta de representatividade da amostra do domínio



15

Explorar as tecnologias de informação e comunicação



16

Potencialidades do e-Assesment

A avaliação eletrónica (*e-Assessment*) permite

- avaliar os mesmos níveis de complexidade cognitiva avaliados na avaliação *paper-based*
 - usar grande variedade de formatos de itens
 - incluir simulações (itens interativos)
 - usar uma grande diversidade de formatos de suportes (recurso à imagem e ao som)
 - a aplicação de testes adaptativos CAT (*Computer Adaptive Testing*), com elevado potencial diagnóstico e formativo
-

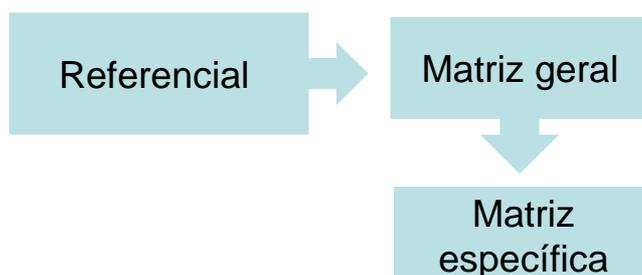
17

Conceptualização e planeamento de um teste

Dos documentos orientadores ao referencial

18

Referencial, matriz e estrutura



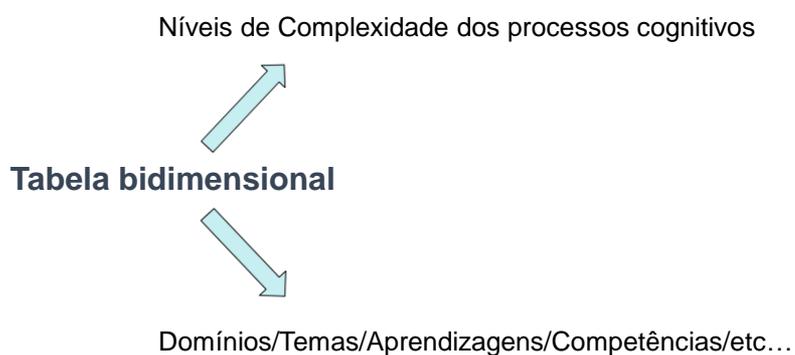
19

Referencial

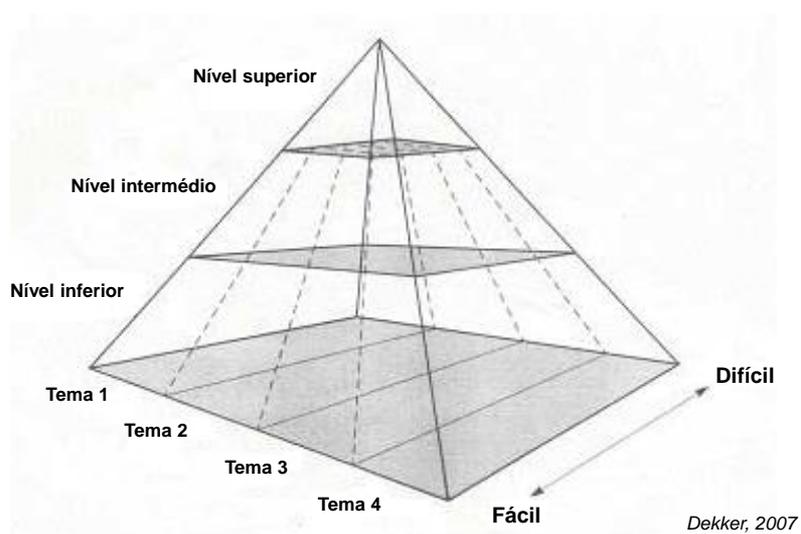
Documentos orientadores: PASEO, Aprendizagens Essenciais, Programas, *Quadro Europeu Comum de Referência...*

Domínios/Temas/Aprendizagens/Competências e Processos Cognitivos que podem ser objeto de avaliação no teste e respetiva relevância curricular, **tendo em conta a necessidade de delimitar o objeto de avaliação.**

20



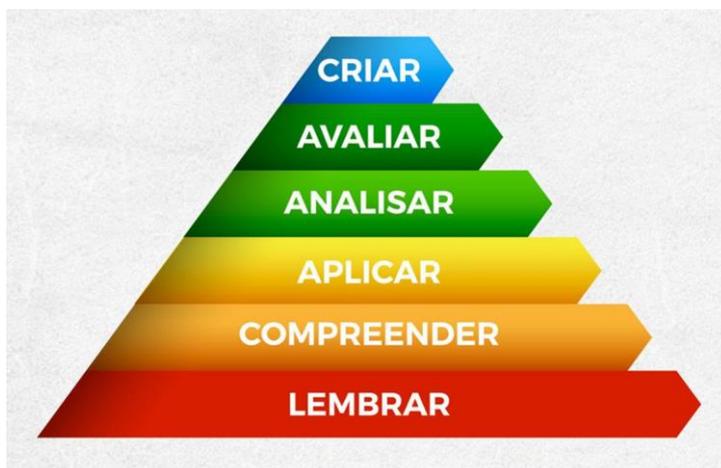
21

Desenho conceptual de uma matriz


22

Níveis de complexidade cognitiva

Taxonomia de Bloom (revista)



Retirado de <https://images.app.goo.gl/DC172sVd2NWrWMCQ7>

23

Níveis de complexidade cognitiva

Taxonomia de Marzano

The Three Systems and Knowledge

Self-System		
Beliefs About the Importance of Knowledge	Beliefs about Efficacy	Emotions Associated with Knowledge

Metacognitive System			
Specifying Learning Goals	Monitoring the Execution of Knowledge	Monitoring Clarity	Monitoring Accuracy

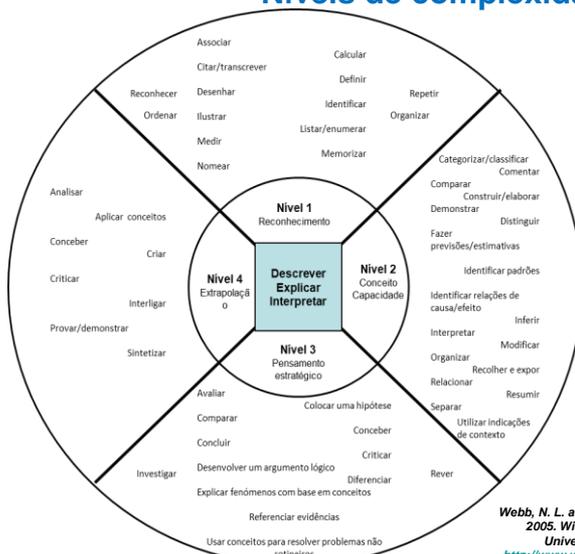
Cognitive System			
Knowledge Retrieval	Comprehension	Analysis	Knowledge Utilization
Recall Execution	Synthesis Representation	Matching Classifying Error Analysis Generalizing Specifying	Decision Making Problem Solving Experimental Inquiry Investigation

Knowledge Domain		
Information	Mental Procedures	Physical Procedures

Robert Marzano (retirado de <https://developingcreativelearnersjack.wordpress.com/>)

24

Níveis de complexidade cognitiva



Webb, N. L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <http://www.wcer.wisc.edu/WAT/index.aspx> (adaptado)

25

Níveis de complexidade cognitiva

ATIVIDADES DE NÍVEL 1	ATIVIDADES DE NÍVEL 2	ATIVIDADES DE NÍVEL 3	ATIVIDADES DE NÍVEL 4
Reconhecer/relembrar elementos da estrutura de uma história (sequência de acontecimentos, personagens, enredo, tempo e espaço).	Identificar ou resumir os acontecimentos principais de uma narrativa.	Justificar ideias com base em pormenores e exemplos.	Implementar um projeto que requeira a especificação de um problema, a conceção e a condução de uma experiência, a análise dos dados e a comunicação de resultados/soluções.
Efetuar cálculos matemáticos elementares.	Usar indicações de contexto para identificar o significado de palavras desconhecidas.	Usar a voz de forma apropriada ao objetivo e ao público.	Aplicar um modelo matemático para clarificar um problema ou uma situação.
Localizar num mapa.	Resolver problemas de rotina com múltiplas etapas.	Identificar questões de investigação e conceber pesquisas para um problema científico.	Analisar e sintetizar informação de múltiplas fontes.
Representar um conceito ou uma relação científica por palavras ou num diagrama.	Descrever a relação de causa/efeito num acontecimento específico.	Desenvolver um modelo científico para uma situação complexa.	Descrever e ilustrar como temas comuns podem ser transversais em diferentes culturas.
Executar procedimentos de rotina, tais como medir um comprimento ou usar os sinais de pontuação corretamente.	Identificar padrões de comportamento.	Determinar o objetivo do autor e descrever o modo como o mesmo pode afetar a interpretação de uma seleção de textos.	Conceber um modelo matemático para informar e resolver uma situação prática ou abstrata.
Descrever as características de um local ou de uma pessoa.	Formular um problema de rotina a partir de dados e condições previamente fornecidos.	Aplicar um conceito a outros contextos.	
	Organizar, representar e interpretar dados.		

Webb, N. L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <http://www.wcer.wisc.edu/WAT/index.aspx> (adaptado)

26

Matriz geral

Nível de complexidade cognitiva Domínios/temas	Nível Inferior	Nível Médio	Nível Superior	Peso (%)
1				35% a 42%
2				25% a 35%
3 ...				25% a 35%
Peso em intervalo(%)	30% a 42%	30% a 42%	25% a 30%	100%

27

Matriz geral

Nível de complexidade cognitiva Domínios	Nível 1		Nível 2		Nível 3	Peso intervalo (%)	
	Identificar informação explícita	Fazer inferências simples	Mobilizar conhecimentos linguísticos em contexto (identificar, aplicar)	Compreender informação implícita	Interpretar e relacionar ideias e informação		Avaliar, criticar e argumentar (ideias, linguagem e elementos textuais). Compor textos.
Oralidade	X	X		X		12% a 18%	
Leitura	X	X		X	X	15% a 25%	
Educação literária		X		X	X	20% a 35%	
Gramática			X		X	15% a 25%	
Escrita					X	25% a 35%	
Peso intervalo (%)	10% a 20%	10% a 20%	10% a 20%	20% a 30%	5% a 15%	20% a 30%	100%

28

Matriz específica – 1º momento

Nível de complexidade cognitiva Domínios	Nível 1		Nível 2			Nível 3	Peso intervalo (%)
	Identificar informação explícita	Fazer inferências simples	Mobilizar conhecimentos linguísticos em contexto (aplicar)	Compreender informação implícita	Interpretar e relacionar ideias e informação	Avaliar, criticar e argumentar (ideias, linguagem e elementos textuais). Compor textos	
Oralidade	2	1	...	1			8% a 15%
Leitura	1	1	1	15% a 25%
Educação literária		2		2	2		20% a 35%
Gramática	1	...	1		15% a 25%
Escrita (parâmetros)				1	35% a 45%
Peso (intervalo da matriz geral) %	28% a 33%		35% a 42%			25% a 30%	100%

29

Matriz específica – 2º momento

Enquadra-se os itens na matriz e apura-se o peso final

Nível de complexidade cognitiva Domínios	Nível 1		Nível 2 -			Nível 3	Peso intervalo (%)
	Identificar informação explícita	Fazer inferências simples	Mobilizar conhecimentos linguísticos em contexto (aplicar)	Compreender informação implícita	Interpretar e relacionar ideias e informação	Avaliar, criticar e argumentar (ideias, linguagem e elementos textuais). Compor textos	
Oralidade	...	1.2 (RR) 10	...	1.1 (EM) 10			14%
Leitura	1.2 (EM) 10	1.3 (ord) 10	18%
Educação literária		4 (RR) 10		2 (EM) 10	3 (RR) 10		22%
Gramática	4.1 (EM) 10	...	5.2 (ass) 8		16%
Escrita (parâmetros)				11 (RE) 44	40%
Peso (%)	18%	15%	22%	18%	5%	22%	100%
Peso (intervalo da matriz geral) %	28% a 33%		35% a 42%			20% a 30%	100%

30

O que nos diz a matriz específica do teste

- Número total de itens do teste
 - Peso / número de itens de cada domínio / tema que o teste avalia
 - Peso / número de itens em cada nível de complexidade cognitiva
 - Níveis de complexidade cognitiva requeridos para responder aos itens do teste
 - Equilíbrio do teste, articulando todos estes aspetos
-

31

Trabalho prático da sessão assíncrona

- **Analisar os documentos orientadores.**
 - **Construir uma matriz de um teste.**
-

32

- Brookhart, S. (2013). *How to create and use rubrics*. Alexandria (USA): ASCD
- Dekker, T. (2007). A model for constructing higher level classroom assessments. *Mathematics Teacher*, 101(1), 56-61.
- Figari, G. (1996). *Avaliar: que referencial*. Porto: Porto Editora
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. The Falmer Press: London.
- Madaus, G., Haney, W. & Kreitzer, A. (2002). The role of testing in evaluations. In D. Stufflebeam, G. Madaus e T. Kellaghan (Eds), *Evaluation models: Viewpoints on educational and human services evaluation*, pp 113-125. Kluwer Academic Publishers: NY.
- Miller, M., Linn, R. & Gronlund, N. (2009). *Measurement and Assessment in Teaching* (10th edition). New Jersey: Pearson
- Peralta, H. (2002). Como avaliar competências. Algumas considerações. In Paulo Abrantes & Filomena Araújo (coords) *Avaliação das Aprendizagens – Das concepções às práticas*. DEB-ME. pp. 25-34
- Ribeiro, L. (1991). *Avaliação da aprendizagem*. 8^a ed. Lisboa: Texto Editora.
- Stufflebeam, D., & Shinkfield, A. (2007). *Evaluation Theory, Models & Applications*. S. Francisco: Jossey-Bass.
- Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. EUA: Jossey-Bass.
-