

## CONSTRUÇÃO, ADAPTAÇÃO E VALIDAÇÃO DA *ESCALA DE AUTO-EFICÁCIA ACADÉMICA* (EAEA)<sup>1</sup>

*Silvia Pina Neves*<sup>2</sup>

*Lúisa Faria*<sup>3</sup>

**Resumo:** As expectativas de auto-eficácia, de acordo com Bandura (2001), referem-se a domínios de realização específicos, pelo que a sua avaliação deve ser microanalítica, o que implica a construção de instrumentos de avaliação adaptados às respectivas especificidades.

Assim, neste artigo, começamos por descrever as várias fases da construção da *Escala de Auto-Eficácia Académica* (EAEA), a saber: (i) definição do seu racional teórico-prático, partindo de revisões bibliográficas e de entrevistas realizadas com alunos e professores; (ii) definição das suas dimensões e redacção dos seus itens; e (iii) pré-teste, que inclui a revisão dos itens por especialistas, reflexões faladas com alunos e um estudo-piloto com 207 alunos.

Finalmente, apresentamos os resultados do estudo de adaptação e validação com 1.302 alunos, constatando-se que as dimensões da EAEA são consistentes, mas pouco discriminativas entre si, e que análises factoriais confirmatórias revelam que um modelo com três factores correlacionados, aceitando-se a existência de covariância entre os erros de alguns dos itens da escala, é o que melhor se ajusta aos dados.

**Palavras-chave:** auto-eficácia académica, qualidades psicométricas, análise factorial confirmatória.

**Construction, adaptation, and validation of the *Academic Self-Efficacy Scale* (ASES) (Abstract):** Self-efficacy expectancies, according to Bandura (2001), are related to particular domains of performance, thus, self-efficacy evaluations must be micro-

---

<sup>1</sup> Este trabalho foi financiado pela Fundação para a Ciência e a Tecnologia, no âmbito da bolsa de doutoramento da primeira autora (*Quadro Comunitário de Apoio III: Programa Operacional Ciência, Tecnologia, Inovação – Medida IV. 3*).

<sup>2</sup> Doutoranda em Psicologia na Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto (FPCE-UP). [stpneves@portugalmail.pt](mailto:stpneves@portugalmail.pt)

<sup>3</sup> Professora Associada com Agregação da FPCE-UP. [lfaria@fpce.up.pt](mailto:lfaria@fpce.up.pt)  
Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto, Rua Dr. Manuel Pereira da Silva, s/n, 4200-392 Porto.

analytic, creating the need to construct specific instruments adapted to the particularities of each domain.

Therefore, we begin this article by describing the various steps in the construction of the *Academic Self-Efficacy Scale* (ASES): (i) the definition of its theoretical and practical background, based on literature reviews and on interviews with students and teachers; (ii) the definition of its dimensions and the conception of its items; and (iii) the pre-test, that includes the revision of the items by experts and its discussion in aloud reflections with students, followed by a preliminary study with 207 students.

Finally, we present the results of an adaptation and validation study with 1.302 students, which evidence the good internal consistency, but the poor discriminative validity of ASES's dimensions, and the confirmatory factor analyses reveal that the model with three correlated factors, allowing the existence of covariance between some items' residuals, fits better to the data.

**Keywords:** academic self-efficacy, psychometric qualities, confirmatory factor analysis.

## Introdução

A auto-eficácia académica pode ser definida como o conjunto de crenças e de expectativas acerca das capacidades pessoais para realizar actividades e tarefas, para concretizar objectivos e para alcançar resultados, no domínio particular da realização escolar (Pina Neves & Faria, 2004). Esta definição tem como referência a abordagem microanalítica (Bandura, 1977, 1997, 2001), a qual pressupõe que as expectativas de auto-eficácia são dependentes dos contextos e das situações de realização concretas, remetendo para as especificidades de cada domínio particular de realização, de cada situação e, até mesmo, de cada tarefa. Por conseguinte, a sua avaliação deve ser circunstancial e contextualizada (Pina Neves & Faria, 2004), implicando, por conseguinte, a construção de instrumentos particulares, capazes de avaliar as expectativas de eficácia que se formam especificamente para os domínios de realização a estudar.

A este propósito, a literatura propõe várias recomendações metodológicas para a construção de instrumentos de avaliação da auto-eficácia, clarificando, por exemplo, quais os cuidados a ter na delimitação dos domínios específicos a avaliar, na redacção e refinamento dos itens dos instrumentos e na escolha do tipo de escala de resposta (Bandura, 2001; Lee & Bobko, 1994; Maurer & Pierce, 1998; Pajares, 1997; Pajares, Hartley & Valiante, 2001; Pina Neves & Faria, 2004). Para além disso, a posterior validação dos instrumentos requer estudos da validade, da fidelidade e da sensibilidade

dos seus itens e dimensões (Ciochină & Faria, 2006; Clark & Watson, 1995; Faria, 1998; Messick, 1995; Pina Neves & Faria, 2004). A este nível, a análise factorial confirmatória dá um importante contributo, pois permite testar e, eventualmente, validar em que medida o modelo teórico subjacente ao racional de um instrumento se ajusta aos dados empíricos, servindo como suporte da validade de constructo do próprio instrumento (Floyd & Widaman, 1995; Maia, 1996; Stapleton, 1997; Thompson & Daniel, 1996).

## Construção da EAEA

A construção da EAEA, conforme se pode verificar resumidamente no Quadro 1, foi estruturada em três fases, comportando cada uma delas três tipos de actividades nucleares, que apresentamos, mais detalhadamente, a seguir.

Quadro 1 – Fases e actividades para a construção da EAEA

Fases	Actividades
1.ª Fase – Fundamentação e definição do racional teórico-prático do instrumento	<ul style="list-style-type: none"> <li>▪ Pesquisas bibliográficas e revisão da literatura</li> <li>▪ Revisão de outros instrumentos de avaliação da auto-eficácia</li> <li>▪ Realização de entrevistas com alunos e com professores</li> </ul>
	<ul style="list-style-type: none"> <li>▪ Definição das dimensões teóricas</li> <li>▪ Redacção das instruções e dos itens</li> <li>▪ Formatação gráfica do instrumento</li> </ul>
2.ª Fase – Concepção e formatação do instrumento	<ul style="list-style-type: none"> <li>▪ Revisão do instrumento por um painel de especialistas</li> </ul>
3.ª Fase – Pré-teste ao instrumento	<ul style="list-style-type: none"> <li>▪ Realização de sessões de reflexão falada com alunos</li> <li>▪ Realização de um estudo-piloto</li> </ul>

Ora, considerando todo este conjunto de recomendações, bem como a necessidade de dispormos de um instrumento que permitisse avaliar as expectativas de eficácia pessoal no contexto particular da realização escolar, foi construída a *Escala de Auto-Eficácia Académica* (EAEA), cujas fases da construção passamos a descrever, sendo, posteriormente, apresentados os resultados do estudo da sua validação preliminar, para o contexto escolar português.

Na 1.<sup>a</sup> Fase – *Fundamentação e definição do racional teórico-prático do instrumento* – pretendíamos constituir um corpo teórico e um referencial prático de conhecimentos que permitissem fundamentar todas as opções conceptuais e metodológicas, feitas ao longo do processo de construção do próprio instrumento. Nesse sentido, foram realizadas pesquisas bibliográficas, privilegiando as produções científicas de autores especialistas no domínio da auto-eficácia, bem como as suas recomendações metodológicas para a construção de instrumentos de avaliação deste constructo. Foram também revistos alguns instrumentos, já existentes, de avaliação da auto-eficácia. Para além disso, foram ainda realizadas entrevistas exploratórias com 24 alunos (12 do 9.º ano e 12 do 10.º ano) e com 8 professores (2 de Língua Portuguesa do 9.º ano, 2 de Matemática do 9.º ano, 2 de Português do 10.º ano e 2 de Matemática do 10.º ano), com o objectivo de explorar as percepções e crenças de eficácia pessoal no domínio escolar, bem como os principais aspectos da sua formação, diferenciação e relação com a realização escolar. Por fim, refira-se que a análise destas entrevistas permitiu organizar um conjunto de pistas para a conceptualização do instrumento de avaliação da auto-eficácia académica e para a posterior redacção dos seus itens.

Na 2.<sup>a</sup> Fase – *Concepção e formatação do instrumento* –, foram definidas as características do instrumento a construir, nomeadamente os seus objectivos de avaliação, as suas dimensões teóricas, o seu formato e a sua escala de resposta, e foram, ainda, redigidos os respectivos itens. Assim, sendo o objectivo do instrumento ‘avaliar as expectativas de auto-eficácia académica para a realização escolar em geral e para a realização em dois domínios mais específicos: a Língua Portuguesa e a Matemática’, foi definido que o instrumento seria uma *escala* com três subescalas que corresponderiam às suas três dimensões teóricas: *auto-eficácia escolar geral*, *auto-eficácia em Português* e *auto-eficácia em Matemática*. A escala de resposta seria de tipo *Likert* com 6 pontos (desde 1 – *Discordo Totalmente* – até 6 – *Concordo Totalmente*), avaliando o nível de concordância dos alunos com os aspectos apresentados em cada item, de tal modo que pontuações mais altas indicariam expectativas de eficácia pessoal mais elevadas. Após este trabalho de conceptualização, procedemos à redacção das instruções e dos itens da escala, redacção esta que foi fundamentada no conjunto de informações teórico-práticas organizado durante a 1.<sup>a</sup> Fase, tendo resultado um conjunto de 28 itens, dos quais 9 se reportam à *auto-eficácia escolar geral*,

8 à *auto-eficácia em Português* e 11 à *auto-eficácia em Matemática*. A escala foi, então, formatada e, posteriormente, designada como *Escala de Auto-Eficácia Académica* (EAEA).

Na 3.<sup>a</sup> Fase – *Pré-teste ao instrumento* –, tivemos como objectivos: (i) testar a acessibilidade do vocabulário utilizado na escala; (ii) garantir a compreensão unívoca das suas instruções e dos seus itens; (iii) testar a adequabilidade e a pertinência dos itens; e (iv) refinar a escala, de modo a conseguir uma versão adequada para administração num estudo-piloto. Deste modo, a EAEA começou por ser revista por um painel de especialistas em Metodologia e Avaliação Psicológica, tendo surgido importantes contributos para a melhoria do conteúdo dos itens, no sentido de uma maior uniformização da sua redacção. Foi também sugerida a eliminação de alguns itens, devido à sua redundância. Assim, após a introdução das sugestões dos especialistas, a EAEA ficou com 26 itens: 8 para a *auto-eficácia escolar geral*, 8 para a *auto-eficácia em Português* e 10 para a *auto-eficácia em Matemática*. Posteriormente, foram realizadas sessões, individuais ou em pequeno grupo, com 12 alunos (6 do 9.º ano e 6 do 10.º ano) de ambos os sexos. Estas sessões incluíram um momento inicial de esclarecimento dos objectivos da sessão, um momento de administração experimental da EAEA e um momento final de reflexão falada com os alunos sobre a escala e o seu preenchimento. Após a realização destas sessões, foram introduzidas pequenas alterações ao nível da redacção de alguns itens da escala. Finalmente, foi realizado um estudo-piloto, no qual a EAEA foi administrada colectivamente a uma amostra de 207 alunos (54,1% de raparigas e 45,9% de rapazes), com idades compreendidas entre os 14 e os 18 anos ( $M = 15,0$ ;  $DP = 0,68$ ), não se tendo observado dificuldades no seu preenchimento.

Quanto aos resultados deste estudo-piloto, ao nível das qualidades psicométricas, verificou-se que os valores de *alpha* de Cronbach encontrados foram superiores a 0,92, revelando uma elevada consistência interna para as três dimensões teóricas, e os coeficientes de assimetria e de curtose revelaram valores próximos dos da distribuição normal.

Por sua vez, uma análise factorial exploratória (análise factorial em componentes principais, com rotação *varimax*) extraiu 4 factores com valores próprios superiores a 1, que explicam 79,9% da variância total dos resultados: os três primeiros factores correspondem às três dimensões teóricas (F1 – *auto-eficácia em Matemática*; F2 – *auto-eficácia em Português*; F3 – *auto-eficácia escolar geral*), surgindo o último factor (F4) como um factor residual, saturado apenas pelos itens n.º 16 e n.º 22 (ambos da *auto-eficácia escolar geral*), o que parece sugerir que estes itens estarão pouco associados à sua dimensão teórica de pertença. Na verdade, estes foram os itens que mostraram funcionar pior no contexto global da EAEA, apresentando uma baixa validade interna e um fraco poder discriminativo.

Assim sendo, considerámos que, sendo este um estudo-piloto e dado o reduzido número de elementos da amostra, seria precoce retirar ou alterar qualquer um dos itens da escala, mostrando-se, para tal, necessária a realização de outros estudos que explorassem estes e outros indicadores de validade, fidelidade e sensibilidade junto de amostras mais alargadas e mais diversificadas, assumindo-se que o processo de adaptação e validação da EAEA será, como não podia deixar de ser, um processo dinâmico.

### Estudo de adaptação e validação da EAEA

#### Método

#### Amostra

Neste estudo participaram 1.302 alunos (48,6% de raparigas e 51,4% de rapazes), com idades compreendidas entre os 13 e os 21 anos ( $M = 15,2$ ;  $DP = 1,05$ ). Destes alunos, 657 frequentavam o 9.º ano (50,5%) e 645 frequentavam o 10.º ano (49,5%) em escolas públicas do Grande Porto. Quanto ao nível socioeconómico (NSE), 30,3% dos alunos pertencia ao NSE Baixo, 36,2% ao NSE Médio e 33,6% ao NSE Alto.

#### Instrumentos e Procedimento

A EAEA foi administrada colectivamente, em tempo e sala de aula, juntamente com um *Questionário Socio-demográfico* também por nós construído. As sessões de administração foram uniformizadas para todos os grupos-turmas.

No que respeita aos tratamentos dos dados e análises estatísticas, utilizámos o *SPSS 12.0 for Windows*, para realizar a análise factorial exploratória, o estudo da consistência interna das dimensões e da validade interna dos itens, a análise da distribuição dos resultados, o poder discriminativo dos itens e as correlações entre dimensões. Utilizámos ainda o *EQS 6.1 for Windows* para realizar as análises factoriais confirmatórias.

Finalmente, é de referir que, quando é considerada a dimensão de *auto-eficácia em Matemática*, o número de sujeitos em análise é de 1.205, em virtude de 97 dos alunos do 10.º ano não frequentarem a disciplina de Matemática.

#### Resultados

##### Análise factorial exploratória (AFE) e consistência interna

À semelhança do estudo-piloto, a análise factorial exploratória extrai 4 factores, com base no critério dos valores próprios superiores a 1 (Tabachnick & Fidell, 2001). A solução obtida, após rotação *varimax*, explica 67,3% da variância total (Quadro 2).

Quadro 2 – Análise factorial em componentes principais, após rotação *varimax*, e índices de validade interna para os itens

Itens		F1	F2	F3	F4	$h^2$	$r^{**}$
03. Vou ter uma boa nota à Matemática no final do ano.	M	0,86				0,82	0,89
15. Vou conseguir resolver exercícios de Matemática, mesmo os que têm cálculos complexos.	M	0,86				0,78	0,84
23. Nos testes de Matemática vou conseguir bons resultados nos exercícios sobre funções.	M	0,85				0,78	0,82
12. Mesmo sabendo que há matérias difíceis, este ano a minha nota a Matemática vai ser boa.	M	0,85				0,82	0,87
26. Vou conseguir melhorar as minhas notas nos testes de Matemática ao longo do ano.	M	0,83				0,78	0,80
21. Nos testes de Matemática vou ter bons resultados nos exercícios sobre geometria.	M	0,79				0,67	0,78
06. As minhas notas nos testes de Matemática vão ser todas positivas.	M	0,79				0,75	0,80
09. A minha nota a Matemática será uma das minhas melhores notas este ano.	M	0,76				0,70	0,76
24. Vou ter bons resultados nos exercícios de Matemática em que se pode usar a calculadora.	M	0,75				0,67	0,75
18. Nos testes de Matemática vou ter bons resultados nos exercícios sobre estatística.	M	0,72				0,57	0,69
07. Vou passar de ano sem ter negativa a nenhuma disciplina.	G	0,54				0,62	0,66
11. Este ano a minha nota a Português vai ser boa, mesmo sabendo que há matérias difíceis.	P		0,80			0,73	0,77
08. Este ano a minha nota a Português vai ser uma das minhas melhores notas.	P		0,80			0,67	0,64
02. No fim do ano vou tirar uma boa nota a Português.	P		0,75			0,69	0,74
05. Vou passar de ano sem ter nenhuma negativa nos testes de Português.	P		0,69			0,60	0,63
20. Nos testes de Português vou ter bons resultados nas perguntas de interpretação.	P		0,62			0,57	0,62
17. Nos testes de Português vou ter bons resultados nas perguntas de gramática.	P		0,59			0,45	0,56
14. Vou conseguir ler e compreender as obras de leitura obrigatória para Português.	P		0,54			0,55	0,61
01. Este ano lectivo vou ser bem sucedido(a) na Escola.	G			0,63		0,64	0,68
10. No final deste ano vou conseguir ter bons resultados na maior parte das disciplinas.	G			0,62		0,73	0,76
13. Este ano vou conseguir tirar boas notas, mesmo nas disciplinas mais difíceis.	G			0,61		0,75	0,80
04. Este ano vou passar com boas notas a todas as disciplinas.	G			0,61		0,75	0,79
16. Vou tirar melhores notas nas disciplinas de que gosto mais.	G		0,58			0,43	0,32
19. Vou conseguir boas notas, mesmo nas disciplinas que têm menos interesse para mim.	G		0,48			0,61	0,71
25. Ao longo do ano vou conseguir melhorar as minhas notas nos testes de Português.	P				0,66	0,71	0,57
22. Vou conseguir melhorar as minhas notas mais baixas ao longo do ano.	G				0,66	0,66	0,50
		Valores Próprios	7,9	4,8	3,0	1,9	
		% Variância Total Explicada	30,2	18,6	11,3	7,2	$\Sigma = 67,3$

Legenda: M – Auto-Eficácia em Matemática; G – Auto-Eficácia Escolar Geral; P – Auto-Eficácia em Português;  $h^2$  – Comunalidades;  $r$  – Índice de Validade Interna dos Itens;  $** p < 0,01$ .

Os coeficientes de saturação e os valores de comunalidade são bons para todos os itens, sendo quase todos superiores a 0,50. Apenas o coeficiente de saturação do item n.º 19 (da *auto-eficácia escolar geral*) e os valores de comunalidade dos itens n.º 16 e n.º 17 (da *auto-eficácia escolar geral* e da *auto-eficácia em Português*, respectivamente) são inferiores.

No que respeita à organização factorial dos itens, os três primeiros factores correspondem às três dimensões teóricas da EAEA: o factor 1 é saturado por todos os itens da *auto-eficácia em Matemática*, em conjunto com o item n.º 7 da *auto-eficácia escolar geral*, sugerindo que a não obtenção de notas negativas no fim do ano lectivo está associada às expectativas de eficácia que os alunos formam especificamente para a disciplina de Matemática; o factor 2 reúne 7 itens da *auto-eficácia em Português*; e o factor 3 reúne 6 itens da *auto-eficácia escolar geral*. Finalmente, o factor 4 é saturado por apenas dois itens, o item n.º 25 da *auto-eficácia em Português* e o item n.º 22 da *auto-eficácia escolar geral*, itens estes que se referem às expectativas dos alunos para melhorar as suas notas ao longo do ano lectivo.

Apesar de esta estrutura factorial isolar claramente os itens de cada uma das três dimensões da EAEA, apoiando a validade discriminativa da escala, as três dimensões estão positivamente correlacionadas entre si, observando-se uma forte correlação entre a *auto-eficácia escolar geral* e as outras duas dimensões, estando estas, por sua vez, moderadamente correlacionadas entre si (Quadro 3). Estas evidências sugerem que as três dimensões da auto-eficácia académica não são independentes, convergindo na avaliação que fazem deste constructo. No mesmo sentido, verifica-se que a média das correlações inter-item é de 0,41, mostrando associações significativas entre os itens das três dimensões da escala. Curiosamente, o item n.º 16 revela correlações fracas com todos os outros itens, mesmo com os da respectiva dimensão teórica de pertença.

Por sua vez, a análise dos valores de *alpha* de Cronbach revela uma boa consistência interna para as três dimensões da EAEA (valores iguais ou superiores a 0,87), sendo a *auto-eficácia em Matemática* aquela que é mais consistente (Quadro 3).

Também a análise dos índices de validade interna dos itens apoia a consistência interna das dimensões, observando-se que a maioria dos itens apresenta uma correlação forte com a dimensão teórica a que pertence (Quadro 2). Os índices mais elevados surgem para os itens da *auto-eficácia em Matemática*, o que sugere que estes avaliam de um modo homogéneo as expectativas de eficácia para o domínio matemático, reforçando a elevada consistência interna dessa dimensão. Mais uma vez, o item n.º 16 é aquele que revela a pior validade interna.

Quadro 3 – Valores de *alpha* de Cronbach para as dimensões de auto-eficácia académica e correlações entre as dimensões

	N.º de itens	<i>alpha</i>	Correlações	
			2.	3.
1. Auto-Eficácia Escolar Geral	8	0,88	0,69 **	0,69 **
2. Auto-Eficácia em Português	8	0,87	–	0,40 **
3. Auto-Eficácia em Matemática	10	0,95	–	–

Nota: \*\*  $p < 0,01$ .

#### Sensibilidade das dimensões e poder discriminativo dos itens

A análise da sensibilidade foi feita para cada dimensão teórica separadamente, verificando-se que os resultados apresentam uma distribuição próxima da normal nas três dimensões (Quadro 4). Os valores das médias, das modas e das medianas estão relativamente próximos, há um bom afastamento entre os valores mínimos e máximos e os coeficientes de assimetria e de curtose são inferiores à unidade.

Quadro 4 – Medidas de tendência central, de dispersão e de distribuição para as dimensões de auto-eficácia académica

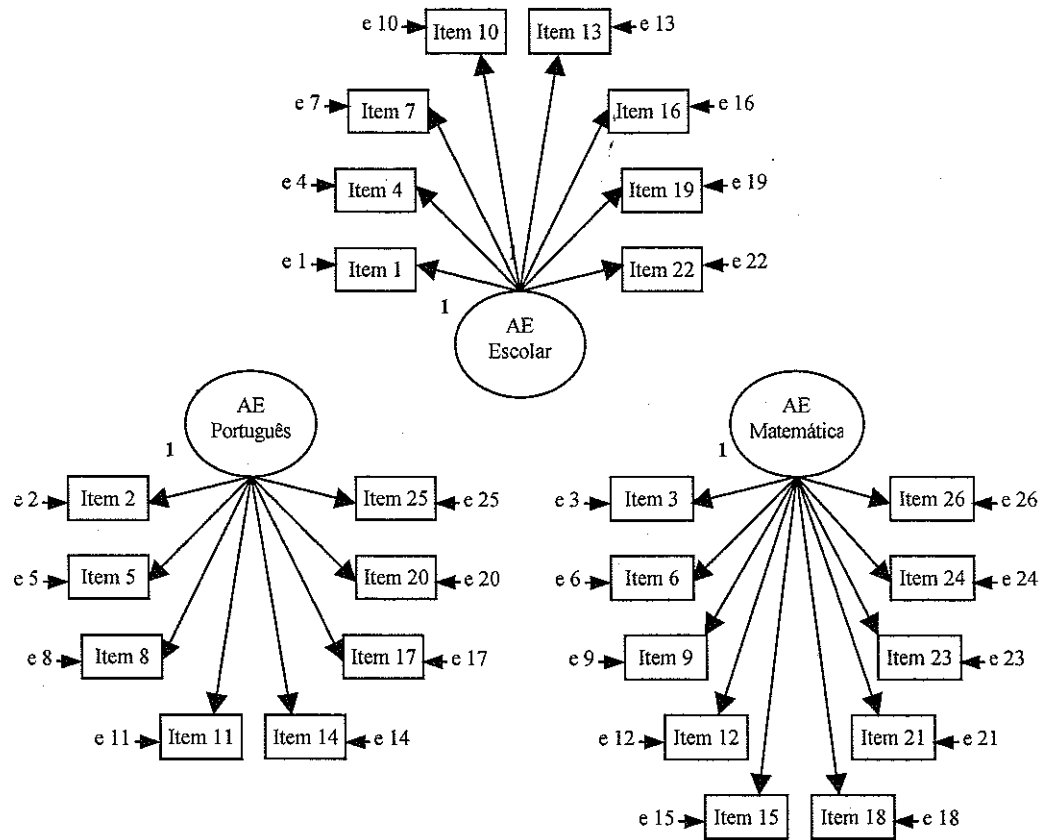
	<i>M</i>	<i>Mo</i>	<i>Md</i>	<i>DP</i>	<i>Mín.</i>	<i>Máx.</i>	<i>Assimetria</i>	<i>Curtose</i>
Auto-Eficácia Escolar Geral	36,2	40	40,0	8,2	13	48	-0,60	-0,27
Auto-Eficácia em Português	34,5	40	36,0	8,1	8	48	-0,65	0,10
Auto-Eficácia em Matemática	41,5	50	45,0	13,6	10	60	-0,69	-0,43

Quanto ao poder discriminativo dos itens, verifica-se que há um domínio na escolha das alternativas de concordância, sobretudo nos itens da *auto-eficácia escolar geral* e da *auto-eficácia em Português*, sugerindo que os alunos têm expectativas de eficácia positivas nestes domínios. Refira-se, ainda, que o item n.º 16 é o que tem pior poder discriminativo, pois concentra mais de 90% das respostas nas 3 alternativas de concordância.

#### Análises factoriais confirmatórias (AFC)

A realização de análises factoriais confirmatórias permitiu testar em que medida o modelo teórico subjacente ao racional da EAEA se ajusta aos nossos dados empíricos. Todas as análises foram realizadas com matrizes de

covariância, construídas automaticamente pelo EQS, a partir da base de dados originais.

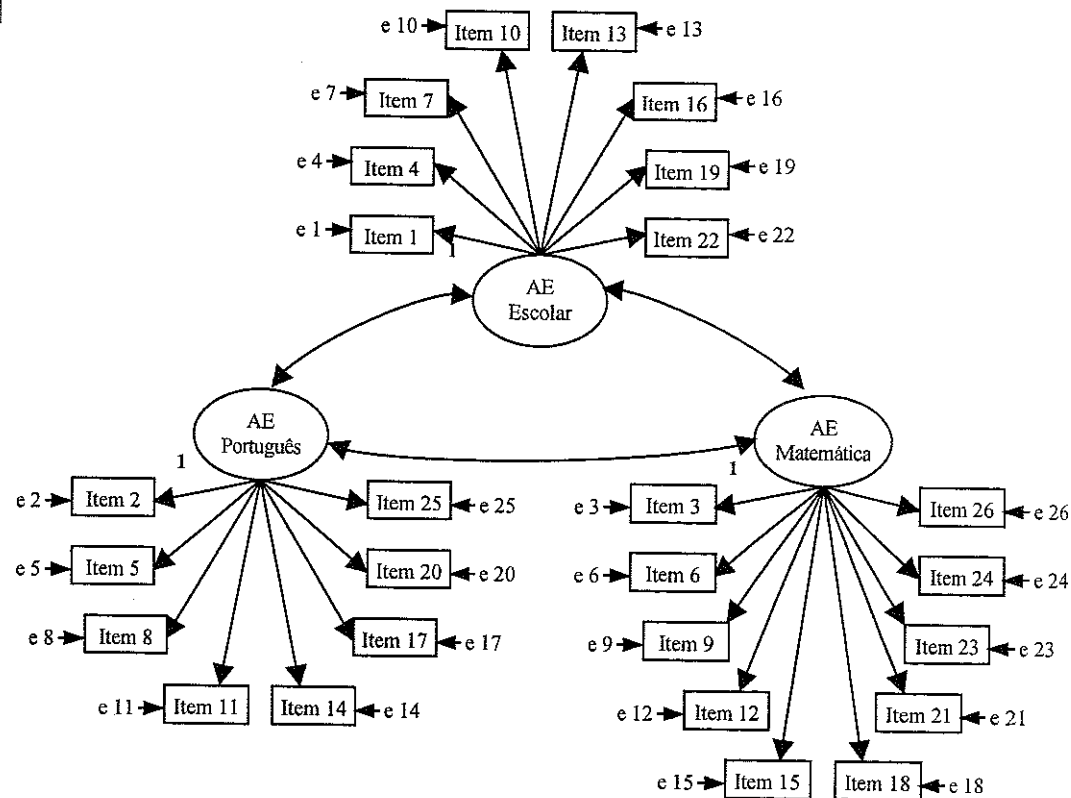


Nota: Todos os parâmetros foram estimados livremente, à excepção da saturação do 1.º item de cada factor, fixada sempre em 1.

Figura 1 – Representação do modelo teórico com três factores não correlacionados (Modelo 1)

Inicialmente, foram testados dois modelos, que encontram fundamento na literatura e na investigação sobre o constructo de auto-eficácia académica. O *Modelo 1* (Figura 1) pressupõe a não existência de correlação entre as três dimensões teóricas da EAEA, fundamentando-se na conceptualização da auto-eficácia como um constructo microanalítico, em que expectativas de eficácia elevadas para um determinado domínio não correspondem,

necessariamente, a expectativas elevadas para outro domínio (Bandura, 1977; Pajares & Miller, 1995). Para além disso, grande parte das investigações que utilizam a AFC apresentam sempre um modelo com factores não correlacionados como o primeiro modelo a ser testado (Bong, 1997; Marsh, Byrne & Shavelson, 1988).



Nota: Todos os parâmetros foram estimados livremente, à excepção da saturação do 1.º item de cada factor, fixada sempre em 1.

Figura 2 – Representação do modelo teórico com três factores correlacionados (Modelo 2)

Por sua vez, o *Modelo 2* (Figura 2) pressupõe a existência de correlação entre as três dimensões teóricas, sendo fundamentado pelas associações significativas que a investigação tem encontrado entre as expectativas de auto-eficácia que se formam para diferentes domínios de realização, nomeadamente no contexto escolar (Bong, 1997, 2004; Pajares, 1996), evidências

estas que também são encontradas no presente estudo, revelando que as três dimensões da EAEA estão correlacionadas entre si.

Refira-se, a propósito, que nestas análises foram considerados índices de ajustamento global<sup>4</sup> e de ajustamento local dos modelos, bem como os valores de  $R^2$ <sup>5</sup> para as respectivas equações estruturais.

Ora, o ajustamento global de ambos os modelos é claramente mau, pois apenas o indicador *RMSEA* está dentro dos valores aceitáveis, mas somente para o modelo que considera a correlação entre factores. A diminuição do valor de  $\chi^2$  no *Modelo 2* é significativa ( $\Delta \chi^2_{(3)} = 1569,7$ ;  $p < 0,001$ ), mas o mau ajustamento revelado pelos restantes indicadores conduz à sua rejeição (Quadro 5).

Quadro 5 – Índices de ajustamento global dos modelos analisados

Modelos	$\chi^2$	gl	$\Delta \chi^2$	$\chi^2 / gl$	NNFI	CFI	RMR <sub>st</sub>	RMSEA
<b>M 0</b> Modelo nulo	23699,2	325	–	–	–	–	–	–
<b>M 1</b> 3 factores não correlacionados	4727,6 *	299	–	15,8	0,79	0,81	0,29	0,11
<b>M 2</b> 3 factores correlacionados	3157,9 *	296	1569,7	10,7	0,87	0,88	0,06	0,09
<b>M 3</b> M 2, com erros correlacionados	2472,1 *	291	685,8	8,5	0,90	0,91	0,06	0,08

Nota: Os modelos analisados são modelos hierárquicos e são apresentados por ordem decrescente de restrição. O índice  $\Delta \chi^2$  permite analisar as diferenças no valor de  $\chi^2$  ao longo do processo de reespecificação.

Legenda:  $\chi^2$  – Qui-Quadrado; NNFI – Bentler-Bonnett Nonnormed Fit Index; CFI – Comparative Fit Index; RMR<sub>st</sub> – Root Mean-Squared Residuals (standardized); RMSEA – Root Mean-Squared Error of Approximation. \*  $\chi^2$  estatisticamente significativo para  $p < 0,001$ .

Em termos de ajustamento local, as soluções standardizadas destes modelos são similares. Apresentam saturações boas para todos os itens, excepto para o item n.º 16 ( $\lambda = 0,33$ , em ambos os modelos), o qual tem também uma grande quantidade de erro associada (*Modelo 1*:  $e = 0,94$ ; *Modelo 2*:  $e = 0,95$ ). Verificamos, ainda, que o factor *auto-eficácia em Matemática* é aquele que globalmente apresenta melhores saturações para os seus itens e erros mais pequenos. No *Modelo 2*, as correlações entre os factores são fortes, sobretudo entre o factor *auto-eficácia escolar geral* e as duas dimensões específicas (Figura 3). Por sua vez, a análise dos coeficientes  $R^2$  (Quadro 6) revela que as equações que menos contribuem para a explicação da variância dos resultados são as que contêm os itens n.º 16, n.º 17, n.º 22 e n.º 25, apresentando sempre um  $R^2 < 0,40$ , em ambos os modelos.

<sup>4</sup> O estudo do ajustamento global dos modelos aos dados empíricos, no âmbito da análise de equações estruturais (*SEM – Structural Equation Modeling*), da qual a análise factorial confirmatória faz parte, deve ser feito com base na análise de diversos indicadores (Boomsma, 2000; Byrne, 1994; Maia, 1996; Yuan & Bentler, 2004). Neste estudo, utilizamos os seguintes índices: o  $\chi^2$  (*Qui-Quadrado*), que avalia a discrepância entre o modelo analisado e a matriz de covariância dos dados (para além disso, quando se analisam modelos hierárquicos, é frequente calcular-se a diferença entre os valores de  $\chi^2$  dos modelos ( $\Delta \chi^2$ ), podendo esta diferença ser sujeita a um teste de significância) (Yuan & Bentler, 2004); o NNFI (*Bentler-Bonnett Nonnormed Fit Index* – equivalente ao *Tucker-Lewis Index* utilizado no LISREL), que compara o modelo analisado com o modelo nulo, sendo adequado para a comparação de modelos hierárquicos, nomeadamente quando o processo de reespecificação conduz à introdução de novos parâmetros nos modelos (Marsh, Balla & McDonald, 1988) – este índice é influenciado pelo tamanho da amostra quando esta é pequena (Boomsma, 2000); o CFI (*Comparative Fit Index*), que também compara o modelo proposto pelo investigador com o modelo nulo e que é igualmente indicado para a comparação de modelos hierárquicos, oferecendo a vantagem de já ter em consideração a dimensão da amostra e de ser relativamente independente desta (Bentler, 1990b, in Byrne, 1994); a RMR<sub>st</sub> (*Root Mean-Squared Residuals – Standardized*), que representa a diferença entre a matriz de covariância original e a matriz ajustada, informando sobre a magnitude dos resíduos do modelo (idealmente, esta diferença deveria ser igual a zero) (Boomsma, 2000); e a RMSEA (*Root Mean-Squared Error of Approximation*), que analisa a aproximação do ajustamento do modelo analisado ao modelo populacional, possuindo um intervalo de confiança (Curran, Bollen, Chen, Paxton & Kyrby, 2003). Para estes indicadores, os valores que permitem aceitar um ajustamento dos modelos aos dados empíricos são:  $\chi^2$  com  $p$  estatisticamente não significativo ou, no caso de este ser significativo,  $\chi^2 / gl \leq 5$  (o cálculo deste rácio tenta corrigir a sensibilidade do  $\chi^2$  à dimensão da amostra, permitindo que se aceite o ajustamento do modelo se  $\chi^2 / gl \leq 5$ , ainda que  $\chi^2$  seja estatisticamente significativo); NNFI > 0,90; CFI > 0,90; RMR<sub>st</sub> < 0,05; e RMSEA < 0,10.

<sup>5</sup> O  $R^2$  (*Squared Multiple Correlation Coefficient*) é um índice que mostra o ajustamento entre o modelo estudado e cada uma das equações da solução standardizada (Boomsma, 2000), representando também a quantidade de variância explicada por cada uma dessas equações (Jöreskog, 2000). Os valores de  $R^2$  variam entre 0 e 1, sendo que valores mais altos indicam uma maior quantidade de variância explicada.

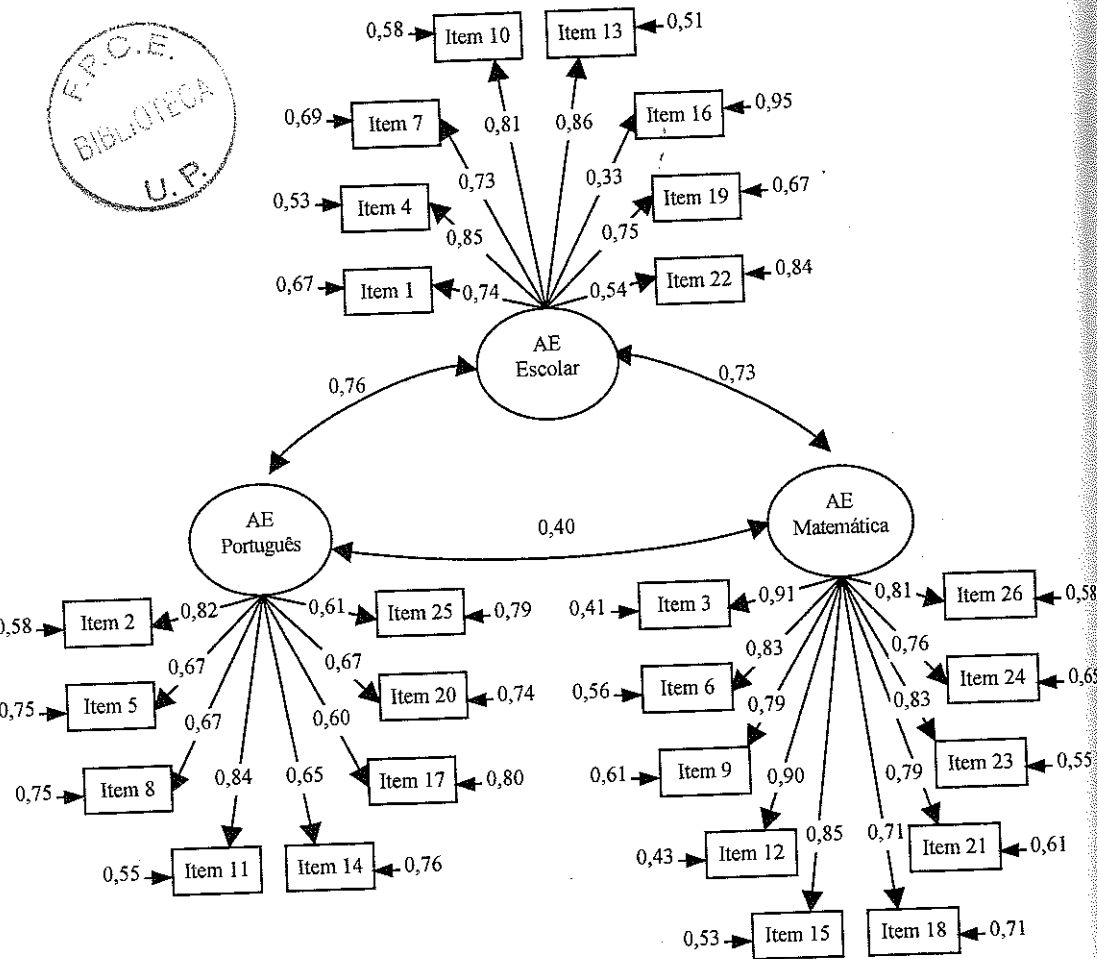


Figura 3 – Representação da solução estandardizada para o Modelo 2

Perante o mau ajustamento global evidenciado anteriormente, procedemos a uma análise dos parâmetros do *Modelo 2* (o modelo com valor de  $\chi^2$  mais baixo), para identificar eventuais parâmetros desajustados (*malfitting parameters*), com vista à reespecificação do modelo. A reespecificação dos modelos, nomeadamente através da introdução de associações entre as variâncias-erro dos itens, é um dos procedimentos mais frequentemente utilizados para promover o seu ajustamento global (Byrne, 2001; Gerbing & Anderson, 1984), uma vez que a introdução de determinados

parâmetros nos modelos contribui, quase sempre, para uma diminuição significativa do valor de  $\chi^2$ . De qualquer modo, é imprescindível que a reespecificação seja fundamentada *a priori* e *a posteriori*, quer estatística quer conceptualmente (Boomsma, 2000; Byrne, 1994; Gerbing & Anderson, 1984): “One major pitfall is when researchers allow error variances to correlate, solely as a tempting means to improve fit, without providing an empirical or theoretical rationale” (Boomsma, 2000: 475).

Quadro 6 – Valores de  $R^2$  para as equações estruturais dos modelos 1, 2 e 3

Equações	Modelo 1	Modelo 2	Modelo 3
v1 = λ F1 + e1	0,55	0,55	0,55
v2 = λ F2 + e2	0,66	0,67	0,67
v3 = λ F3 + e3	0,83	0,83	0,84
v4 = λ F1 + e4	0,73	0,72	0,72
v5 = λ F2 + e5	0,43	0,45	0,43
v6 = λ F3 + e6	0,68	0,69	0,65
v7 = λ F1 + e7	0,50	0,53	0,48
v8 = λ F2 + e8	0,46	0,45	0,44
v9 = λ F3 + e9	0,62	0,62	0,63
v10 = λ F1 + e10	0,67	0,66	0,66
v11 = λ F2 + e11	0,70	0,70	0,70
v12 = λ F3 + e12	0,81	0,82	0,82
v13 = λ F1 + e13	0,75	0,74	0,74
v14 = λ F2 + e14	0,42	0,43	0,43
v15 = λ F3 + e15	0,73	0,72	0,72
v16 = λ F1 + e16	0,11	0,11	0,11
v17 = λ F2 + e17	0,35	0,36	0,36
v18 = λ F3 + e18	0,50	0,50	0,49
v19 = λ F1 + e19	0,55	0,56	0,56
v20 = λ F2 + e20	0,46	0,45	0,45
v21 = λ F3 + e21	0,63	0,63	0,63
v22 = λ F1 + e22	0,28	0,29	0,30
v23 = λ F3 + e23	0,70	0,69	0,69
v24 = λ F3 + e24	0,58	0,58	0,58
v25 = λ F2 + e25	0,38	0,37	0,40
v26 = λ F3 + e26	0,67	0,66	0,67

Legenda:  $R^2$  – Squared Multiple Correlation Coefficient.

Concretizando, no plano estatístico, é recomendado o cálculo da diferença no valor de  $\chi^2$  ( $\Delta\chi^2$ ), a par da utilização de outros indicadores de ajustamento global (Yuan & Bentler, 2004), bem como a análise da matriz residual dos itens, procurando as magnitudes de residuo mais altas (*largest*



*standardized residuals* – Byrne, 2001). Já no plano conceptual, é importante não só analisar, do ponto de vista semântico, os itens envolvidos no processo de reespecificação, mas também a congruência dos novos parâmetros com a teoria subjacente ao modelo.

Assim, a análise dos parâmetros foi feita com base nas sugestões do *LM Test* (*Lagrange Multiplier Test – Multivariate Statistics*), para as covariâncias-erro, tendo sido identificados cinco parâmetros que, ao serem introduzidos, levariam a uma diminuição significativa do valor de  $\chi^2$ , a saber: (e7, e5), (e25, e22), (e26, e25), (e7, e6) e (e6, e5), por ordem de apresentação pelo *LM Test*. Portanto, a reespecificação do *Modelo 2* parece justificar-se, quer estatística, quer conceptualmente: do ponto de vista estatístico, os erros dos parâmetros propostos referem-se aos pares de itens que, na matriz residual estandardizada, apresentam as magnitudes de resíduo mais altas (*largest standardized residuals* – Byrne, 2001), sugerindo que os erros associados a estes pares de itens podem estar a afectar o ajustamento global do modelo, enquanto que, do ponto de vista conceptual, os cinco parâmetros em questão representam covariâncias-erro entre itens que partilham algo em comum, nomeadamente os pares de itens n.ºs 7 e 5, n.ºs 7 e 6 e n.ºs 6 e 5 (referem a expectativa para transitar de ano sem notas negativas ou com notas todas positivas) e os pares de itens n.ºs 25 e 22 e n.ºs 26 e 25 (referem a expectativa para melhorar as piores notas ao longo do ano). Ou seja, assim, parece justificar-se a reespecificação do modelo, pelo que procedemos à introdução dos cinco parâmetros<sup>6</sup>, originando um modelo reespecificado – *Modelo 3* (Quadro 5).

Deste modo, a propósito do ajustamento global deste modelo reespecificado, a primeira observação recai sobre a diminuição significativa do valor de  $\chi^2$  ( $\Delta \chi^2_{(5)} = 685,8$ ;  $p < 0,001$ ), bem como sobre o aumento dos valores de *NNFI* e de *CFI*, que revelam agora um ajustamento global aceitável. Por sua vez, o valor de *RMR<sub>st</sub>* mantém-se e o de *RMSEA* melhora ligeiramente. Ora, apesar destas evidências serem um forte argumento para a inclusão dos cinco parâmetros no modelo com três factores correlacionados, é ainda necessário observar a solução estandardizada para o *Modelo 3* (Figura 4), analisando particularmente os valores das covariâncias-erro introduzidas: como podemos ver, estas são todas significativas, revelando correlações de magnitude moderada entre os respectivos erros (correlações entre 0,30 e 0,44). Por sua vez, o ajustamento local dos itens é semelhante ao do *Modelo 2*, não se observando alterações significativas nas correlações entre factores, nem nos coeficientes  $R^2$  (Quadro 6).

<sup>6</sup> A reespecificação do modelo foi feita passo a passo (*single-step modification*), tendo sido introduzidos os parâmetros um a um e analisadas as respectivas consequências, ao nível do ajustamento global e local. No entanto, por questões de simplificação, só são apresentados os resultados do modelo reespecificado após a introdução dos cinco parâmetros.

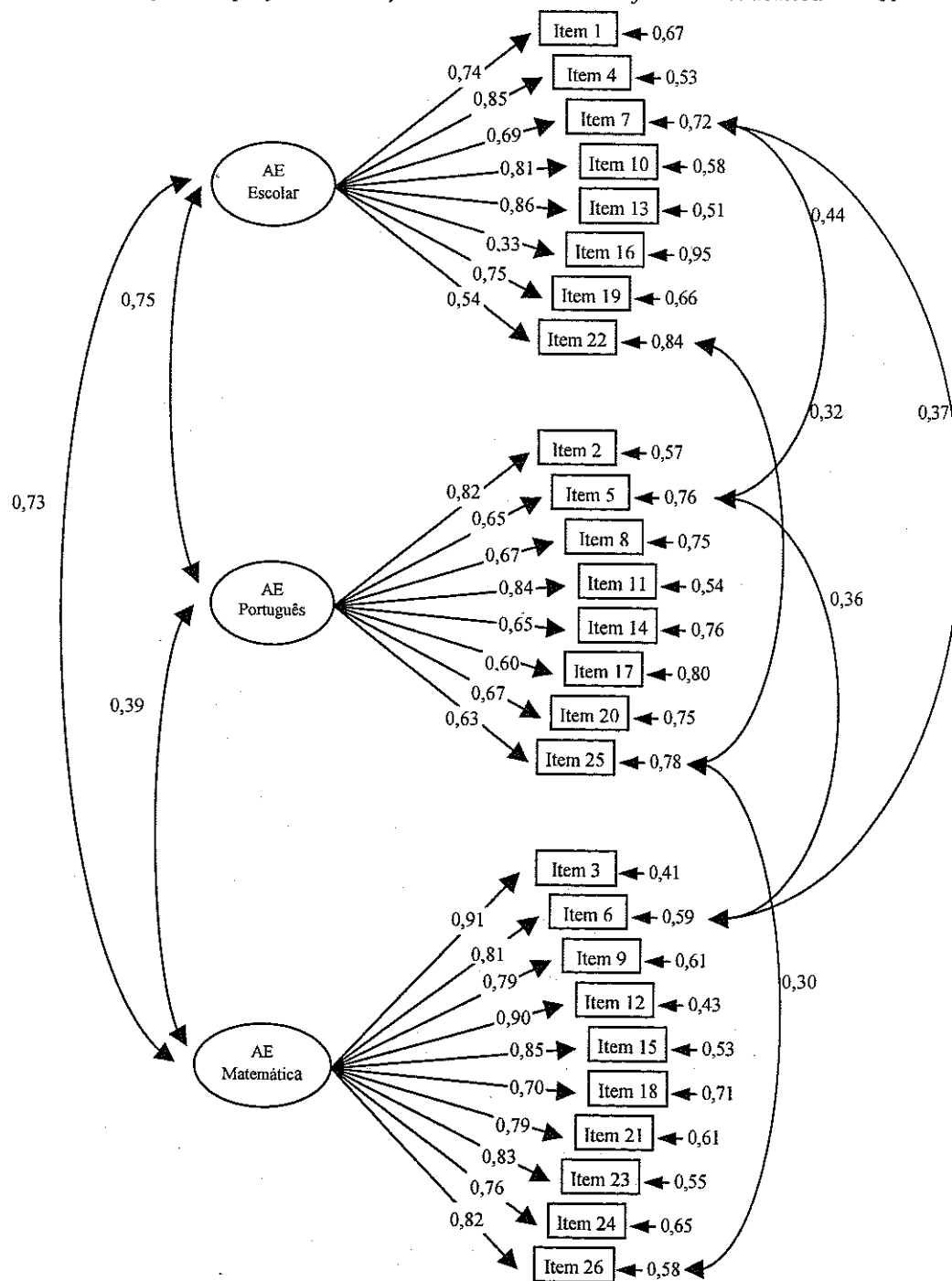


Figura 4 – Representação da solução estandardizada para o Modelo 3

Em suma, assim sendo, estes resultados, considerados em conjunto com a melhoria que a reespecificação produz ao nível do ajustamento global do modelo, parecem apoiar a definição de um modelo de auto-eficácia académica que considere que as suas três dimensões estão correlacionadas e que inclua ainda covariâncias entre os erros de alguns dos seus itens: as consequências conceptuais e metodológicas deste novo modelo são discutidas a seguir.

### Discussão e conclusões

Neste artigo, foram apresentados os primeiros resultados sobre a consistência interna, a sensibilidade e a validade de constructo da *Escala de Auto-Eficácia Académica*.

A análise factorial exploratória evidenciou a existência de 4 factores, 3 dos quais remetem claramente para as 3 dimensões teóricas da EAEA e, destes 3 factores, a *auto-eficácia em Matemática* é aquele com maior poder explicativo da variância dos resultados (30,2%), sugerindo uma maior importância das expectativas de eficácia para esta disciplina, relativamente às que se formam para o domínio da Língua Portuguesa e da realização escolar em geral.

A nosso ver, este facto deve-se à relevância que a Matemática, enquanto disciplina, assume no contexto e no currículo escolares, nomeadamente em Portugal, onde se registam elevadas taxas de insucesso (GAVE/ME, 2004). A Matemática é, assim, percebida como uma das disciplinas mais difíceis do currículo, sendo responsável pela diferenciação entre “bons” e “maus” alunos e influenciando a realização escolar global (Pina Neves & Faria, submetido a publicação): de facto, a investigação tem concluído que as expectativas de eficácia para o domínio da Matemática são um factor determinante do desempenho escolar (OCDE, 2004). Consequentemente, estas particularidades fazem com que a Matemática se distinga facilmente de todas as outras áreas disciplinares. Talvez, também por isso, a *auto-eficácia em Matemática* seja a dimensão mais importante e a mais consistente, e os seus itens apresentem uma forte validade interna.

Não obstante a estrutura factorial encontrada, que isola os itens de cada uma das três dimensões da EAEA, verificou-se que estas se encontram positivamente correlacionadas entre si, sendo que as correlações entre a *auto-eficácia em Português* e a *auto-eficácia em Matemática* são moderadas, o que se deve, provavelmente, ao facto de estas serem duas dimensões específicas, esperando-se, por conseguinte, menores níveis de associação (Bandura, 1977; Pajares & Miller, 1995). De modo semelhante, Bong (1997, 2004) encontra correlações moderadas entre as expectativas de eficácia para o domínio verbal e para o domínio matemático.

Já as correlações entre a *auto-eficácia escolar geral* e as duas dimensões mais específicas (*Português* e *Matemática*) são elevadas. A associação entre a *auto-eficácia escolar geral* e a *auto-eficácia em Português* pode estar relacionada com o facto de a Língua Portuguesa representar um domínio que é transversal a todo o currículo escolar (Valadares, 2003). De facto, “só desenvolvendo capacidades [ao] nível [da] leitura, compreensão e expressão orais e escritas, em Língua Portuguesa, [é que] os alunos estarão aptos a construir a sua aprendizagem (...) noutras áreas [disciplinares] e a alcançar o sucesso” (Valadares, 2003: 31). Também neste sentido, Carvalho (1999) refere que a expressão escrita é um factor importante no sucesso escolar em geral, ideia que se torna ainda mais pertinente se considerarmos que os testes escritos são o principal meio de avaliação dos alunos. Por sua vez, a associação entre a *auto-eficácia escolar geral* e a *auto-eficácia em Matemática* pode estar relacionada com o papel preponderante que a disciplina de Matemática parece ter no currículo escolar e no próprio desempenho académico, o que foi já aqui discutido. Para além disso, há evidências que mostram que a relação com o rendimento escolar tende a ser mais forte quando se analisam as crenças de auto-eficácia matemática do que as crenças de eficácia para a realização noutras áreas escolares (Pajares, 1996), o que sugere que as percepções de competência pessoal, para o contexto de realização escolar, podem, antes de mais, passar pelas percepções de mestria e de eficácia no domínio da Matemática.

Ora, ainda que interpretável à luz da literatura e da investigação, a forte associação observada entre a dimensão geral de auto-eficácia académica e as duas dimensões específicas mostra que há convergência na forma como estas avaliam a auto-eficácia – ao contrário do que seria desejável para dimensões que deveriam ser específicas e microanalíticas –, podendo colocar em causa a sua validade discriminativa (Messik, 1995). Por seu turno, também a média das correlações inter-item (0,41) é elevada (Clark & Watson, 1995), se considerarmos que a avaliação da auto-eficácia deve ser dirigida para domínios específicos. Deste modo, ou será desnecessária a existência de uma dimensão geral de auto-eficácia académica, ou, de acordo com Gorsuch (1997) e Gerbing e Anderson (1984), poderemos colocar a hipótese da existência de um factor de 2.<sup>a</sup> ordem, com base na magnitude das correlações observadas entre os 26 itens e entre as três dimensões. A este propósito, refira-se, ainda, que os resultados das análises factoriais confirmatórias parecem também apoiar esta hipótese.

Como vimos, as AFC possibilitaram aprofundar o estudo da validade de constructo da EAEA e do modelo teórico subjacente ao seu racional, ao mesmo tempo que permitiram uma análise mais pormenorizada dos seus itens e do seu contributo individual para a avaliação das expectativas de auto-eficácia académica.

Nestas AFC, verificamos que o modelo com 3 factores correlacionados e com um grupo de erros correlacionados é aquele que revela melhor ajustamento global comparativamente aos outros modelos, apresentando melhores valores ao nível dos indicadores *NNFI*, *CFI*, *RMSEA*, e um poder explicativo de cerca de 91% da covariância dos dados, apesar de o seu valor de  $\chi^2$  ser ainda estatisticamente significativo. A este propósito, saliente-se que é possível haver contradições entre diferentes indicadores estatísticos: por vezes, observam-se valores de  $\chi^2$  significativos (indicando um mau ajustamento do modelo estudado), a par de outros indicadores que revelam um ajustamento aceitável do mesmo modelo, o que pode ser explicado pelo facto de a estatística de  $\chi^2$  ser sensível à dimensão da amostra (Boomsma, 2000; Yuan & Bentler, 2004). De qualquer modo, há uma diminuição significativa do valor de  $\chi^2$  relativamente aos outros modelos, o que, à partida, juntamente com a melhoria observada ao nível dos restantes indicadores, apoia a aceitação deste modelo para explicar a auto-eficácia académica.

Ora, esta melhoria no ajustamento global deve-se, certamente, à introdução de novos parâmetros no modelo analisado, nomeadamente, da covariância entre factores e entre alguns erros, o que faz aumentar a capacidade explicativa desse modelo perante os dados empíricos (Gerbing & Anderson, 1984). Como vimos, a hipótese da covariância entre factores foi empiricamente fundamentada, o mesmo acontecendo para as covariâncias entre os erros. No entanto, a introdução de novos parâmetros nos modelos, muito embora contribua para a melhoria do ajustamento global dos mesmos, tem implicações ao nível conceptual e metodológico que devem ser ponderadas.

Uma das implicações, do *foro conceptual*, é a necessidade de serem feitos ajustes ao nível do racional do instrumento, de tal modo que se considerem associações até aí não previstas, o que só deve ser feito caso essas associações façam sentido, do ponto de vista teórico (Boomsma, 2000). A outra das implicações, do *foro metodológico*, remete-nos para o facto de que a necessidade de introduzir a estimação de um grupo de covariâncias-erro para melhorar o ajustamento do modelo pode ser indicativa da presença de factores de 2.<sup>a</sup> ordem. E isto porque os erros associados aos itens representam algo que é externo aos factores (que não é contemplado por estes), e que pode ser considerado não só como um erro de medida, mas também como uma componente que, não pertencendo a um factor de 1.<sup>a</sup> ordem, pode estar a definir um factor de 2.<sup>a</sup> ordem (Gerbing & Anderson, 1984).

Assim, antes de aceitarmos o primeiro modelo que oferece um bom ajustamento aos dados, devemos, numa atitude mais parcimoniosa, testar outros modelos em função de um trabalho coordenado entre a teoria e as evidências estatísticas (Floyd & Widaman, 1995), modelos esses que podem ser tão bons ou melhores do que o primeiro modelo encontrado (Boomsma, 2000; Gerbing & Anderson, 1984).

Finalmente, a análise dos 26 itens da escala sugere que o item n.º 16 (“*Vou tirar melhores notas nas disciplinas de que gosto mais*”) é o pior item no contexto global da nova escala. Estatisticamente, este item é aquele que apresenta uma maior quantidade de erro associada e pior coeficiente  $R^2$ , o que mostra o seu fraco contributo para a avaliação da auto-eficácia escolar geral. Tem também um mau poder discriminativo e um índice de validade interna muito fraco, o que reforça a ideia de ser um mau indicador para avaliar a auto-eficácia escolar geral. Conceptualmente, este item parece ser um pouco redundante – pois o interesse pelas disciplinas parece estar intimamente associado a um maior investimento por parte dos alunos –, não acrescentando valor crítico ao conjunto de itens da escala e não tendo nada, no seu conteúdo, que promova a sua sensibilidade na avaliação de diferentes níveis de auto-eficácia académica.

Em suma, perante estes resultados, parece-nos importante prosseguir a investigação das qualidades psicométricas da EAEA. Globalmente, as dimensões desta nova escala revelam uma boa consistência interna, mas uma baixa validade discriminativa. Também encontramos evidências que sugerem que alguns dos itens podem não ser bons indicadores para a avaliação da auto-eficácia académica, sendo necessários outros estudos para explorar estes resultados.

Assim, seria importante replicar as análises aqui realizadas com outras amostras. Também as análises factoriais confirmatórias necessitam de ser aprofundadas, propondo outros modelos para a organização dos itens. A este nível pode ser interessante a análise de modelos alternativos, que proponham, por exemplo, uma estrutura hierárquica para a auto-eficácia académica – com a existência de factores de segunda ordem –, uma vez que este tipo de estrutura já demonstrou ser ajustada para a avaliação das expectativas de auto-eficácia no contexto escolar (Bong, 1997, 2004). Também seria útil realizar estudos de invariância métrica – importantes no desenvolvimento de novos instrumentos de avaliação (Byrne & Watkins, 2003; Floyd e Widaman, 1995) –, de forma a testar a estabilidade da estrutura da EAEA entre grupos de sujeitos de ambos os sexos, de diferentes anos de escolaridade/níveis etários e de diferentes níveis socioeconómicos.

Finalmente, saliente-se que consideramos este estudo de adaptação e validação como um estudo preliminar da validade de constructo da EAEA, entendendo que a validação de um instrumento é um processo dinâmico e contínuo (Nunally, 1977, *in* Faria, 1998), que depende das características dos contextos e das amostras em estudo, bem como do próprio factor “tempo”, pelo que serão necessárias mais evidências empíricas para que possamos compreender a adequabilidade dos itens e das dimensões da EAEA e, por conseguinte, eliminar ou reformular os piores itens, no sentido

do refinamento desta nova escala de avaliação das expectativas de auto-eficácia académica.

## Referências

- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84 (2), 191-215.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New-York: Freeman.
- Bandura, A. (2001). *Guide for constructing self-efficacy scales (Revised)*. Available from Frank Pajares, Emory University.
- Bong, M. (1997). Generality of academic self-efficacy judgments: Evidence of hierarchical relations. *Journal of Educational Psychology*, 89 (4), 696-709.
- Bong, M. (2004). Academic motivation in self-efficacy, task value, achievement goal orientations, and attributional beliefs. *The Journal of Educational Research*, 97 (6), 287-297.
- Boomsma, A. (2000). Reporting analysis of covariance structures. *Structural Equation Modeling*, 7 (3), 461-483.
- Byrne, B. M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, application and programming*. Thousand Oaks, CA: Sage Publications.
- Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing*, 1 (1), 55-86.
- Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, 34 (2), 155-175.
- Carvalho, J. A. B. (1999). A expressão escrita, factor de sucesso escolar. In P. F. Pinto (Org.), *Português, propostas para o futuro – Transversalidades* (pp. 89-97). Lisboa: Associação de Professores de Português.
- Ciochină, L., & Faria, L. (2006). Concepções pessoais de inteligência de estudantes portugueses e romenos: Estudo preliminar de análise factorial confirmatória. *Psychologica*, 41, 171-191.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7 (3), 309-319.
- Curran, P. J., Bollen, K. A., Chen, F., Paxton, P., & Kyryby, J. B. (2003). Finite sampling properties of the point estimates and confidence intervals of the RMSEA. *Sociological Methods & Research*, 32 (2), 208-252.
- Faria, L. (1998). *Desenvolvimento diferencial das concepções pessoais de inteligência durante a adolescência*. Lisboa: Fundação Calouste Gulbenkian. Junta Nacional de Investigação Científica e Tecnológica.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7 (3), 286-299.
- GAVE/ME (2004). *PISA 2003 – Resultados do estudo internacional*. Lisboa: Gabinete de Avaliação Educacional do Ministério da Educação.
- Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, 11, 572-580.
- Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of Personality Assessment*, 68 (3), 532-560.
- Jöreskog, K. G. (2000). *Interpretation of R<sup>2</sup> revisited*. Documento disponível em [www.ssicentral.com/lisrel/techdocs/r2rev.pdf](http://www.ssicentral.com/lisrel/techdocs/r2rev.pdf).
- Lee, C., & Bobko, P. (1994). Self-efficacy beliefs: Comparison of five measures. *Journal of Applied Psychology*, 79 (3), 364-369.
- Maia, J. (1996). Um discurso metodológico em torno da validade de constructo: Posições de um Lisrelita. In L. S. Almeida, S. Araújo, M. M. Gonçalves, C. Machado & M. R. Simões (Eds.), *Avaliação Psicológica: Formas e Contextos* (Vol. IV, pp. 43-59). Braga: APPORT.
- Marsh, H. W., Balla, J., & McDonald, R. P. (1988). Goodness of fit in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103 (3), 391-410.
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-construct: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80 (3), 366-380.
- Maurer, T. J., & Pierce, H. R. (1998). A comparison of Likert scale and traditional measures of self-efficacy. *Journal of Applied Psychology*, 83 (2), 324-329.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50 (9), 741-749.
- OCDE (2004). *Learning for tomorrow's world – First results from PISA 2003*. Paris: OCDE.
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66 (4), 543-578.
- Pajares, F. (1997). Current directions in self-efficacy research. In M. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.
- Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: greater discrimination increases prediction. *Measurement & Evaluation in Counseling & Development*, 33, 214-221.
- Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology*, 42 (2), 190-198.
- Pina Neves, S., & Faria, L. (2004). Auto-eficácia académica: Definição conceptual e recomendações metodológicas para a construção de instrumentos de avaliação. In C. Machado, L. S. Almeida, M. Gonçalves & V. Ramalho (Orgs.), *Avaliação Psicológica: Formas e Contextos* (Vol. X, pp. 391-399). Braga: Psiquilíbrios Edições.
- Pina Neves, S., & Faria, L. (submetido a publicação). Auto-eficácia académica e atribuições causais em Português e Matemática. *Análise Psicológica*.

- Stapleton, C. D. (1997). *Basic concepts and procedures of confirmatory factor analysis*. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX, em 24 de Janeiro.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn & Bacon.
- Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement*, 55 (2), 197-208.
- Valadares, L. M. (2003). *Transversalidade da Língua Portuguesa*. Porto: Edições ASA.
- Yuan, K.-H., & Bentler, P. M. (2004). On chi-square difference and z tests in mean and covariance structure analysis when the base model is misspecified. *Educational and Psychological Measurement*, 64 (5), 737-757.