

Uma nova maneira de avaliar as competências escritoras na Redação do ENEM

■ Ruben Klein*

■ Nilma Fontanive**

Resumo

O artigo apresenta um estudo realizado com a correção da redação do ENEM 2008 adotando-se uma nova forma de avaliar o domínio das competências escritoras do exame. Cada competência e seus critérios de correção foram tratados como 5 itens polítomos com ordenação de resposta de 1 a 7. Essa metodologia possibilita que se obtenha para cada competência/item polítomos as estatísticas clássicas, percentuais de resposta por categoria, coeficientes ponto-polisseriais e ponto-bisseriais por categoria de resposta. Permite também o uso da Teoria da Resposta ao Item (TRI) para obter-se uma escala de proficiências de Redação no ENEM. A adoção de metodologia de interpretação de escalas permite descrever as competências e habilidades dominadas pelos candidatos nos diferentes níveis da escala, possibilitando definir os pontos nos quais as categorias de resposta das competências são dominadas. O uso dessa metodologia pode aperfeiçoar a apresentação dos resultados do ENEM e ser útil para maior aprofundamento dos resultados da redação. Com esse trabalho os autores pretendem introduzir uma discussão do uso da TRI na correção da redação do ENEM.

Palavras-chave: ENEM. Correção de redação e avaliação de competências escritoras. Teoria da Resposta ao Item e interpretação de escala.

A new way to evaluate the writing skills in ENEM's compositions

Abstract

This article presents a study based on the correction of ENEM 2008 compositions using a new way of evaluating the students' writing skills. Each skill and its correction criteria was treated as five different items ranging from

* Doutor em Matemática, Massachusetts Institute of Technology, EUA; Consultor da Fundação Cesgranrio.
E-mail: ruben@predialnet.com.br

** Doutora em Educação, PUC-Rio; Coordenadora de Avaliação em Larga Escala da Fundação Cesgranrio.
E-mail: nilmaf@cesgranrio.org.br

5 to 7. This methodology allows to obtain the classical statistics, percentage of response by category and credit coefficient per category for each skill. It also allows the use of the item tri response theory in order to obtain a proficiency scale in ENEM's compositions. The adoption of this methodology allows to describe the candidates skills and abilities in the different levels of the scale, making possible to define the points in which the categories of response skills are dominated. The use of this methodology may improve the presentation of ENEM's results and it may also be useful for further studies of the compositions results. With this work, the authors intend to introduce a discussion on the use of the Tri in the correction of ENEM's compositions.

Keywords: ENEM. Compositions correction and writing skills evaluation. Item Response Theory and Scale interpretation

Un nuevo modo de Evaluar las Competencias Escritoras en la Redacción del ENEM

Resumen

El artículo presenta estudio realizado en la a corrección de la redacción del ENEM 2008 adoptándose nova forma de evaluar el dominio de las competencias escritoras de los candidatos. Cada competencia y sus criterios de corrección fueran tratados como 5 itens polítomos con ordenación de respuesta de 1 a 7. Esa metodología posibilita que se obtenga para cada competencia/item polítomos las estadísticas clásicas, porcentuales de respuesta por categoría, coeficientes punto-poliseriales e punto-biseriales por categoría de respuesta. Permite también el uso de la Teoría de la Respuesta al Ítem-Tri para obtenerse una escala de resultados proficientes con la Redacción del ENEM. La adopción de metodología de interpretación de escalas permite describir las competencias y habilidades dominadas por los candidatos en los diferentes niveles de la escala, posibilitando definir los puntos los cuales las categorías de respuesta de las competencias son dominados. El uso de esa metodología puede perfeccionar la presentación de los resultados del ENEM y ser útil para mayor profundidad de los resultados de la redacción. Con ese trabajo los autores pretenden introducir una discusión del uso de la TRI en la corrección de la redacción del ENEM

Palabras clave: ENEM. Corrección de redacción y evaluación de competencias escritoras. Teoría de la Respuesta al Ítem e Interpretación de escala.

O ENEM: breve contextualização

Criado em 1998, o ENEM contou com um número relativamente pequeno de participantes na sua primeira edição 115,6 mil, e a cada ano, nesses dez anos de existência vem aumentando o número de participantes.

A popularização definitiva do ENEM veio em 2004, quando o Ministério da Educação instituiu o Programa Universidade para Todos (ProUni) e vinculou a concessão de bolsas em IES privadas à nota obtida no Exame. Nos anos seguintes, o ENEM vem aumentando o número de participantes, alcançando na sua última edição, em 2008, a marca histórica de 4.018.050 de inscritos e 2.920.560 presentes no exame.

O principal incentivo para que os concluintes e egressos do ensino médio façam o ENEM é a possibilidade concreta de ingressar no ensino superior, pois, a nota obtida no exame pode significar tanto uma bolsa integral ou parcial do ProUni quanto ajudar na conquista de uma vaga em algumas das mais prestigiadas instituições de ensino superior do País.

Ao completar dez anos, o ENEM ocupa um lugar de destaque na agenda educacional brasileira não só por sua contribuição na democratização do acesso ao ensino superior e, em particular na proposta em discussão no País nesse ano de 2009 de ele se transformar em um exame vestibular nacional, unificando o acesso ao ensino superior em todas as universidades brasileiras.

A avaliação das competências escritoras no ENEM

A prova de redação do ENEM propõe-se a avaliar 5 competências escritoras do candidato. Para tanto, a matriz de competências do ENEM foi adaptada para avaliar o desempenho do participante como produtor de um texto no qual ele demonstre capacidade de reflexão sobre o tema proposto. Essa reflexão faz-se partir da leitura dos textos que compõem a proposta, conjugada à leitura da realidade. Há o compromisso de que os temas abordem questões de ordem política, social, cultural ou científica, desde que apresentados como uma situação-problema, para a qual o autor do texto deverá propor soluções, respeitando os direitos humanos.

Os critérios de avaliação baseiam-se nas cinco competências expressas na matriz do ENEM traduzidas para a situação específica de produção de texto. Cada competência desdobra-se em níveis correspondentes aos conceitos *Insuficiente*, *Regular*, *Bom* e *Excelente*, quantificados da seguinte maneira: nível 1, nota 2,5; nível 2, nota 5,0; nível 3, nota 7,5 e nível 4, nota 10,0.

A nota global da Redação é aferida pela média aritmética simples das notas atribuídas a cada uma das cinco competências e o resultado convertido em escala centesimal.

A redação que não atende à proposta da prova (tema e estrutura) recebe o conceito D – Desconsiderada e a redação absolutamente ilegível também é desconsiderada. Os variados casos de redação desconsiderados recebem a nota zero. Quando é apresentada totalmente em branco ou em até sete linhas escritas, título exclusivo, a redação recebe o conceito B – em branco.

Abaixo serão resumidos os critérios utilizados na correção do texto produzido pelo participante do ENEM.

Critérios de Correção	
	Descrição
Competência I	Demonstrar domínio da norma culta da língua escrita
Critérios de Correção 1:	Demonstra conhecimento "precário" da norma culta: inadequação na escolha da variedade lingüística," Graves e frequentes" desvios gramaticais e transgressões "inaceitáveis" das convenções da escrita. Neste nível o participante emprega expressão muito próxima da modalidade oral e registro inadequado à proposta da prova; demonstra inadequação vocabular; revela conhecimento precário das normas gramaticais e das convenções básicas da escrita.
2:	Demonstra conhecimento razoável da norma culta: problemas na escolha da variedade lingüística; desvios gramaticais e transgressões das convenções da escrita pouco aceitáveis nessa etapa de escolaridade. Considera-se como "pouco aceitável" o uso que constitua desvio da sintaxe de concordância, regência e colocação, da pontuação e da flexão.
3:	Demonstra "bom" domínio da norma culta (ainda que com "pontuais" desvios gramaticais ou transgressões pontuais das convenções da escrita). Espera-se que participante empregue um léxico produtivo, ortografia e pontuação pertinentes, demonstre um bom domínio da sintaxe e das regras de flexão. Neste nível podem ocorrer alguns desvios em relação à norma escrita culta.
4:	Demonstra "muito bom" domínio da norma culta (ainda que com "um ou outro deslize" relativo à norma gramatical ou às convenções da escrita). O nível 4 é atribuído a textos em que o produtor tenha plena autonomia com relação à utilização da norma escrita culta, ainda que possa apresentar um ou outro deslize.

Quadro 1: Competência 1.

Fonte: Os autores (2008).

Critérios de Correção	
	Descrição
Competência II	desenvolve outro tema e/ou elabora outra estrutura
Critérios de Correção 1:	Desenvolve "tangencialmente" o tema em um texto que apresenta características do tipo de texto dissertativo-argumentativo; ou apresenta "embrionariamente" o tipo de texto dissertativo-argumentativo (sem "fugir" do tema). Considera-se tangencial o texto que se aproxima apenas por generalidades do tema proposto, ou o texto resultante da adaptação forçada de um projeto de texto ao tema selecionado para a redação. Alguns casos podem apontar textos preparados previamente à realização da prova. Outro caso de atribuição do nível 1 é a apresentação de uma argumentação mínima (embrionária) inserida num texto de outra natureza (tal como carta e narrativa), diferente da estrutura dissertativa proposta.

Continuação

2:	Desenvolve "razoavelmente" o tema, a partir de considerações próximas do senso comum ou por meio de paráfrases dos textos-estímulo, e domina "precariedade" ou "razoavelmente" o tipo de texto dissertativo-argumentativo. O nível 2 indica a presença de um texto construído com base em considerações próximas do senso comum, com pouca reflexão do participante e muita repetição de chavões e clichês; a estrutura é pouco encadeada e sem progressão. Nesse caso, o participante revela ausência de projeto e demonstra tentar discutir o tema à medida que vai escrevendo.
3:	Desenvolve "bem" o tema, mesmo apresentando argumentos previsíveis, e domina "bem" o tipo de texto dissertativo-argumentativo, com indícios de autoria. O nível 3 indica um bom desenvolvimento do tema, como consequência de sua correta compreensão, ainda que com argumentos previsíveis que não impedem, contudo, o surgimento de indícios de autoria. Evidencia-se um bom domínio das técnicas dissertativo-argumentativas.
4:	Desenvolve "muito bem" o tema, a partir de um projeto pessoal de texto e de um repertório cultural produtivo; domina "muito bem" o tipo de texto dissertativo-argumentativo. Este nível indica o desenvolvimento seguro e autônomo do tema, a partir de um repertório cultural produtor e de considerações que fogem ao senso comum. O domínio da estrutura dissertativa argumentativa é muito bom. O projeto de texto do participante define-se com base na compreensão plena da proposta e é desenvolvido segundo a articulação de diferentes perspectivas para a defesa do ponto de vista. Revela ainda marcas de autoria.

Quadro 2: Competência 2.

Fonte: Os autores (2008).

Critérios de Correção	
	Descrição
Competência III	Selecionar, relacionar, organizar e interpretar informações, fatos, opiniões e argumentos em defesa de um ponto de vista
Critérios de Correção 1:	Apresenta informações, fatos e opiniões "precariedade" relacionados ao tema. O nível 1 indica textos com problemas de coerência decorrentes da maneira como as informações, os fatos e as opiniões relacionam-se entre si ou com o tema; ou problemas de coerência de textos que, explicitamente, procuram adequar ao tema proposto o projeto desenvolvido para outro tema (caso comentado na Competência II, nível 1). Nesse nível não há seleção de argumentos.

Continuação

2:	Apresenta informações, fatos e opiniões "razoavelmente" relacionados a um embrião de projeto de texto e/ou limita-se a reproduzir os elementos fornecidos pela proposta de redação. O nível 2 indica textos que apresentam uma estrutura fragmentada ou circular, com eventuais contradições; também se aplica a textos que se limitam a reproduzir ou a parafrasear os argumentos, fatos e opiniões constantes da proposta de redação.
3:	Seleciona informações, fatos, opiniões e argumentos, "relacionando-os" ao seu projeto de texto. O nível 3 indica textos realizados a partir de um projeto pessoal, com boas seleção e interpretação de informações, fatos, opiniões e argumentos que resultam bem relacionados a esse projeto.
4:	Seleciona, interpreta e organiza informações, fatos, opiniões e argumentos, estabelecendo uma relação produtiva entre essa seleção e seu projeto de texto. O nível 4 indica textos realizados a partir de um projeto pessoal, com excelente seleção, interpretação e organização de informações, fatos, opiniões e argumentos que se relacionam de forma produtora a esse projeto. Observa-se que a seleção de elementos feita pelo participante contribui para a defesa coerente do ponto de vista destacado no seu projeto de texto.

Quadro 3: Competência 3.

Fonte: Os autores (2008).

Critérios de Correção	
	Descrição
Competência IV	Demonstrar conhecimento dos mecanismos linguísticos necessários para a construção da argumentação
Critérios de Correção 1:	Articula "precarientemente" as partes do texto. A estrutura de superfície típica do nível 1 é a justaposição de palavras e frases pouco relacionadas, deixando o estabelecimento da coesão a cargo do leitor.
2:	Articula "razoavelmente" as partes do texto, apresentando "problemas frequentes" na utilização dos recursos coesivos. O texto de nível 2 indica uma grande dificuldade do produtor na utilização da modalidade escrita; mesmo que tente articulá-lo, esbarra em problemas não solucionados durante a escolarização. O participante demonstra não ter conseguido incorporar à sua produção textual as regras que regem a articulação do texto formal.
3:	Articula "bem" as partes do texto (ainda que apresente "problemas pontuais" na utilização dos recursos coesivos). O texto de nível 3 demonstra domínio dos mecanismos de coesão, embora apresente ainda alguns desvios com relação à utilização desses mecanismos.
4:	Articula "muito bem" as partes do texto (ainda que apresente "eventuais deslizes" na utilização de recursos coesivos). O texto típico do nível 4 indica que o produtor tem plena autonomia com relação à utilização dos mecanismos coesivos da norma escrita culta, ainda que possa apresentar um ou outro deslize.

Quadro 4: Competência 4.

Fonte: Os autores (2008).

Critérios de Correção	
	Descrição
Competência V	Elaborar proposta de intervenção para o problema abordado, demonstrando respeito aos direitos humanos
Critérios de Correção 1:	Elabora proposta "precariedade" relacionada ao tema, respeitando os direitos humanos. O nível 1 indica que houve pouca compreensão do tema e que a proposta apresentada apenas o tangencia.
2:	Elabora proposta "razoavelmente" relacionada ao tema, mas "não articulada" ao texto desenvolvido, respeitando os direitos humanos. O nível 2 indica textos que apresentam proposta relacionada ao tema, mas sem vinculação ao texto desenvolvido
3:	Elabora proposta "bem" relacionada ao tema, mas "pouco articulada" à discussão desenvolvida no texto, respeitando os direitos humanos. O participante preocupa-se em elaborar proposta explícita para a problemática discutida, no entanto, ela resulta pouco articulada à discussão.
4:	Elabora proposta "bem" relacionada ao tema e "bem articulada" à discussão desenvolvida no texto, respeitando os direitos humanos. O participante elabora proposta explícita para a situação-problema e articula-a coerentemente à discussão.

Quadro 5: Competência 5.

Fonte: Os autores (2008).

Itens polítomos e suas estatísticas clássicas.

Itens de resposta construída pelos alunos que são corrigidos atribuindo-se graus de correção ordenados, além do certo ou errado são exemplos de itens polítomos com categorias de resposta graduadas.

Esses itens não são muito frequentes nas avaliações em larga escala no Brasil, mas têm ganho destaque no Brasil, por causa do Programa Internacional de Avaliação de Alunos (PISA). Mas esses itens são utilizados amplamente em outros países como nos Estados Unidos e Inglaterra. No Brasil, também, em 1995, o SAEB apresentou nas duas disciplinas avaliadas, alguns itens de resposta construída (KLEIN; FONTANIVE, 1995) e em 1997 apenas em Matemática.

Para podermos utilizar as estatísticas clássicas de itens e a TRI na análise da correção da redação do ENEM, temos que considerar cada competência como um item polítomos ordenado com as gradações de resposta dadas pelos critérios de correção (1, 2, 3 e 4) especificados na seção anterior.

É necessário, entretanto, chamar a atenção para o fato do ENEM utilizar a dupla correção de cada redação, por dois avaliadores independentes, e caso haja discrepância de mais de dois pontos na média das notas de 0 a 10, a redação é corrigida por um terceiro avaliador, também independente. No caso de a discrepância ser de dois pontos ou menor, a nota em cada competência é obtida pela média dos graus

atribuídos pelos dois avaliadores. Cada competência fica assim com 7 graduações (1; 1,5; 2; 2,5; 3; 3,5 e 4). Esse é o motivo pelo qual a análise realizada nesse estudo apresenta 7 graduações e não as 4 previstas nos critérios de correção.

Utilizando os bancos de dados da correção da redação de 2008, com os 2853115 registros dos candidatos, cedidos pelo INEP, onde para candidato tem-se sua nota em cada competência e a nota geral, calcularam-se as proporções das notas atribuídas em cada competência, tabela 1.

As estatísticas obtidas para cada competência/item e suas categorias de resposta do ENEM são apresentadas nas tabelas 1 e 2.

Tabela 1: Proporções de respostas por competência e categoria.

Competência	Categoria de resposta							Soma
	1	2	3	4	5	6	7	
1	0.010	0.054	0.189	0.309	0.314	0.105	0.019	1.000
2	0.024	0.091	0.229	0.305	0.254	0.083	0.015	1.000
3	0.045	0.143	0.301	0.287	0.172	0.045	0.007	1.000
4	0.052	0.153	0.298	0.280	0.165	0.046	0.007	1.000
5	0.050	0.156	0.300	0.284	0.163	0.042	0.006	1.000

Fonte: Os autores (2008).

Como se pode ver na tabela 1, a competência 5, "Elaborar proposta de intervenção para o problema abordado, demonstrando respeito aos direitos humanos" é a mais difícil, para os candidatos, pois as proporções em cada uma das colunas das categorias de 5 a 7 são decrescentes, isto é, em cada uma delas há menor proporção de candidatos na competência 5. Em geral, a ordem de dificuldade, em cada categoria, cresce da competência 1 para a 5.

Tabela 2: Coeficientes ponto-poliserials e coeficientes ponto-biserials por competência/categoria.

Competência	Categoria de resposta							
	Pbis1	Pbis2	Pbis3	Pbis4	Pbis5	Pbis6	Pbis7	ppolis
1	-0.216	-0.351	-0.408	-0.117	0.343	0.395	0.256	0.769
2	-0.313	-0.426	-0.375	0.001	0.405	0.428	0.267	0.854
3	-0.414	-0.467	-0.254	0.206	0.467	0.380	0.206	0.887
4	-0.435	-0.462	-0.223	0.220	0.458	0.378	0.202	0.884
5	-0.401	-0.434	-0.219	0.213	0.442	0.354	0.189	0.835

Fonte: Os autores (2008).

O coeficiente ponto-polisserial é a correlação entre uma variável politômica e a medida global envolvendo todas as variáveis. No caso da redação, essa medida é a nota do candidato e as variáveis são as graduações de resposta. Essa é uma medida de discriminação. Pode-se ver que a menor discriminação é obtida pela competência 1, "Demonstrar domínio da norma culta da língua escrita", e a as maiores, pelas competências 3 e 4. Pode-se ver também que os coeficientes ponto-bisseriais são positivos por competência/categoria nas categorias mais altas e negativos nas mais baixas como esperado, pois lembra-se que o coeficiente ponto-bisserial para uma competência/categoria é proporcional à diferença entre a média da nota dos alunos que obtiveram essa graduação e a média geral de todos os alunos.

Chama-se a atenção de que na graduação 4, os coeficientes ponto-bisseriais são negativos ou quase zero para as competências 1 e 2, enquanto são positivas para as demais. Isso novamente indica que a competência 1 é a mais dominada pelos alunos.

Interpretação dos resultados na escala de 25 a 100

Para essa seção empregou-se a técnica de interpretação de escala utilizada no SAEB e estendida a itens polítomos ordenados. Cada competência/graduação corresponde a um item âncora em um nível na escala de 25 a 100. A nota mínima nas redações válidas (retirando-se as desconsideradas e as em branco) é 25 e corresponde à graduação 1 em todas as competências. Por outro lado, a nota 100 corresponde à graduação 7 (4 nos critérios de correção) em todas as competências.

A análise gerou a tabela 3 indicando, para cada nível a ser interpretado, seus "itens âncora".

Para a descrição da habilidade em cada competência/graduação recorre-se aos critérios de correção apresentados nos quadros de 1 a 5. É importante assinalar que as graduações 1, 3, 5 e 7 correspondem, respectivamente, aos critérios 1, 2, 3 e 4.

Por exemplo, a descrição do nível 75 da escala será:

"Os alunos nesse nível demonstram as habilidades de:

Selecionar informações, fatos, opiniões e argumentos, "relacionando-os" ao seu projeto de texto. O nível 3 indica textos realizados a partir de um projeto pessoal, com boas seleção e interpretação de informações, fatos, opiniões e argumentos que resultam bem-relacionados a esse projeto. (C3, critério de correção 3),

Articular "bem" as partes do texto (ainda que apresente "problemas pontuais" na utilização dos recursos coesivos). O texto de nível 3 demonstra domínio dos mecanismos de coesão, embora apresente ainda alguns desvios com relação à utilização desses mecanismos. (C4, critério de correção 3).

Elaborar proposta "bem" relacionada ao tema, mas "pouco articulada" à discussão desenvolvida no texto, respeitando os direitos humanos. O participante preocupa-se em elaborar proposta explícita para a problemática discutida, no entanto, ela resulta pouco articulada à discussão (C5, critério de correção 3).

Tabela 3: Competência/graduação como "itens âncoras" em seus níveis.

Competência	Graduação	nível
C1	2	31.25
C2	2	31.25
C3	2	37.50
C4	2	37.50
C5	2	37.50
C1	3	37.50
C2	3	43.75
C3	3	50.00
C4	3	50.00
C5	3	50.00
C1	4	50.00
C2	4	56.25
C3	4	62.50
C4	4	62.50
C5	4	62.50
C1	5	68.75
C2	5	68.75
C3	5	75.00
C4	5	75.00
C5	5	75.00
C1	6	81.25
C2	6	81.25
C3	6	87.50
C4	6	87.50
C5	6	87.50
C1	7	93.75
C2	7	93.75
C3	7	93.75
C4	7	93.75
C5	7	100.00

Fonte: Os autores (2008).

É importante acrescentar, que como nas escalas do SAEB, as interpretações são cumulativas, ou seja, as habilidades dominadas pelos alunos em um nível incluem as habilidades descritas nos níveis anteriores.

Desse modo, no exemplo apresentado, os alunos dominam também os critérios 3 de correção das competências 1 e 2, âncoras no nível anterior 68,75.

Deve-se ressaltar ainda ordenação obtida e mostrada na tabela 3, onde aparecem primeiro a ordenação da graduação e dentro de cada graduação, a ordenação das competências de 1 a 5, novamente ressaltando que a competência 1 é a que é dominada primeiro pelos alunos, e a 5, a última.

Utilização da Teoria da Resposta ao Item

A consideração de cada competência como um item politomo ordenado permite a modelagem da probabilidade de um candidato obter uma determinada graduação de uma competência por um modelo da Teoria de Resposta ao Item como o de crédito parcial generalizado ou o graduado de Samejima. Ambos estendem o modelo de dois parâmetros para itens dicotômicos para itens politômicos

A figura abaixo ilustra um exemplo das curvas características de um modelo de crédito parcial com 4 graduações. A curva em preto corresponde à graduação 1 (errado em um item de resposta construída), a curva em vermelho corresponde à graduação 2, a curva em verde a graduação 3 e finalmente a curva em azul a graduação 4 (totalmente certo).

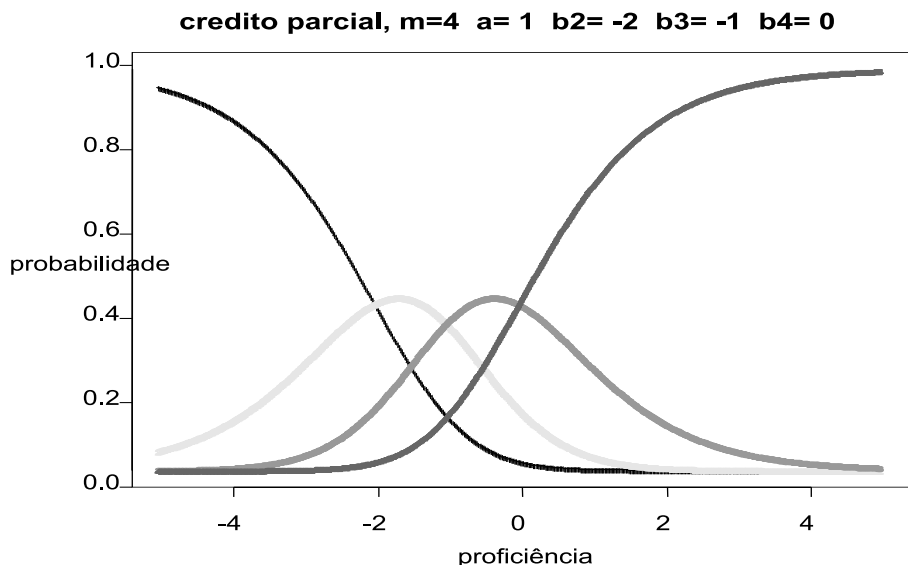


Figura 1: Curvas características de um modelo de crédito parcial com 4 graduações. Fonte: Bock e Muraki (2003).

Utilizou-se o *software* Parscale (BOCK; MURAKI, 2003) para fazer a calibração dos "itens" segundo o modelo acima. Posteriormente foi realizada também uma interpretação na escala de proficiências obtida.

Os parâmetros dos itens obtidos estão na tabela 4. Observa-se que os parâmetros "a", parâmetro da discriminação, variam muito. O menor é o da competência 1, e o maior o da competência 3. De certa maneira, é coerente com os coeficientes ponto-polisseriais anteriormente apresentados na tabela 2.

Tabela 4: Parâmetros estimados das competências (itens politômicos), modo logístico.

Competência	a	b2	b3	b4	b5	b6	b7
1	1.08	-3.16	-2.2	-0.9	0.13	1.76	2.97
2	1.87	-2.39	-1.47	-0.47	0.42	1.6	2.66
3	3.18	-1.9	-1	-0.02	0.83	1.82	2.76
4	2.63	-1.85	-0.95	0.02	0.87	1.87	2.83
5	1.59	-2.04	-1.05	0.04	0.98	2.13	3.12

Fonte: Os autores (2008).

A tabela 5 apresenta os valores de níveis e seus "itens âncoras". O nível está na escala tri da população com média 0 (zero) e desvio padrão 1, enquanto o nível transformado refere-se à população com média 250 e desvio padrão 50.

Observa-se na tabela 5 a mesma ordenação por graduação, mas há algumas diferenças na ordenação das competências dentro das graduações. Por exemplo, na graduação 1, a competência 5 é a 3ª e não a 5ª. Do mesmo modo na graduação 7, a competência 1 é a 4ª e não a 1ª. Isso difere do observado na análise com as notas apresentadas na tabela 3.

Provavelmente essa diferença é devida ao fato da TRI levar em conta além da dificuldade, também a discriminação.

Pode-se ver também na tabela 5, que os níveis correspondentes ao domínio das graduações 3 (critério de correção 2), 5 (critério de correção 3) e 7 (critério de correção 4) de todas as competências são 212,5, 312,5 e 420, cujas diferenças em torno de 100 correspondem a 2 desvios padrões da população.

De fato algo semelhante pode ser observado na tabela 3, pois a média da nota da população é 59,6 e o desvio padrão é 12,9. Os níveis correspondentes aos do parágrafo anterior são: 50, 75 e 100, cujas diferenças de 25 pontos são quase dois desvios padrões.

Conclusões

Esse artigo procura apresentar uma maneira diferente de interpretar as escalas das notas atribuídas à redação do ENEM, aplicando a metodologia de interpretação de escalas para itens polítomos. Essa interpretação foi feita para a escala das notas e uma escala de proficiências obtidas pela TRI. Os autores pretendem também com esse trabalho introduzir a TRI na análise de redações como as do ENEM. Um próximo passo será procurar uma forma de equalizar os resultados obtidos nas redações em diferentes anos.

Tabela 5: Competência/graduação como "itens âncoras" em seus níveis.

Competência	Graduação	nível	nível transformado
C1	2	-2.75	112.5
C2	2	-2.00	150.0
C5	2	-1.75	162.5
C3	2	-1.50	175.0
C4	2	-1.50	175.0
C1	3	-1.50	175.0
C2	3	-1.00	200.0
C3	3	-0.75	212.5
C4	3	-0.75	212.5
C5	3	-0.75	212.5
C1	4	-0.50	225.0
C2	4	-0.25	237.5
C3	4	0.25	262.5
C4	4	0.25	262.5
C5	4	0.50	275.0
C1	5	0.50	275.0
C2	5	0.75	287.5
C3	5	1.00	300.0
C4	5	1.00	300.0
C5	5	1.25	312.5
C2	6	1.75	337.5
C3	6	2.00	350.0
C4	6	2.00	350.0
C1	6	2.25	362.5
C5	6	2.25	362.5
C3	7	2.75	387.5
C2	7	3.00	400.0
C4	7	3.00	400.0
C1	7	3.50	420.0
C5	7	3.50	420.0

Fonte: Os autores (2008).

Referências

BAKER, F. B. *Item Response Theory: parameter estimation techniques*. New York: Marcel Dekker, 1992.

BOCK, D.; MURAKI, E. *Parscale irt scaling, item analysis, and scoring or rating scale data*. Lincolnwood, IL: SSI, 2003.

CESPE; FUNDAÇÃO CESGRANRIO. *Relatório do ENEM 2008*. Rio de Janeiro, 2008. v. 1.

KLEIN, R. Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB). *Ensaio: avaliação e políticas públicas em educação*, Rio de Janeiro, v. 11, n. 40, p. 283-296, jul./set. 2003.

KLEIN, R.; FONTANIVE, N. Avaliação em larga escala. *Em Aberto*, Brasília, DF, v. 15, n. 66, p. 29-34, 1995.

Recebido em: 29/09/2009

Aceito para publicação em: 30/10/2009