

TESTAGEM DE PROFICIÊNCIA EM LEITURA EM INGLÊS: A VISÃO DOS EXAMINANDOS IMPORTA?*

Hamilton de Godoy WIELEWICKI (*Unijuí/RS*)

ABSTRACT: This paper aims at discussing the particular situation of ESP proficiency testing at Universidade Federal de Santa Maria (UFSM). The main argument presented is that the interpretation of final scores in ESP testing should take into account the inter-relationship between the test itself, the self-evaluation of the testees about their reading ability in ESP, and their performance on the test, in order to approach the testing situation from a wider perspective.

0. Introdução

Devido à importância conferida à habilidade de leitura em inglês no contexto acadêmico brasileiro, a exigência da comprovação de tal habilidade, normalmente através de testes de proficiência, já faz parte do cotidiano dos cursos de mestrado e doutorado do país. Apesar de tal importância, ainda parece haver uma carência de estudos especificamente abordando questões relativas à testagem de proficiência em leitura em inglês no contexto brasileiro. Dentre essas questões, poder-se-ia destacar o acentuado índice de reprovação dos examinandos em tais testes observado, por exemplo, na UFSM (ver Wielewicki, 1997b).

Esse fato, se encarado simplisticamente sob a ótica do princípio de seletividade intelectual preconizado no Parecer 977/65 (que regulamenta a pós-graduação no Brasil), indicaria que tais examinandos não estão plenamente aptos a ingressarem em um curso de mestrado. Ocorre, no entanto, que esses mesmos examinandos já são, em sua maioria, estudantes dos cursos de mestrado da UFSM e, muitos deles, em estágio de defesa de dissertação. Assim, a situação que se estabelece é paradoxal. De um lado, a justificativa para tais testes é de que sejam capazes de identificar o nível adequado de habilidade de leitura em ESP com o qual um estudante de mestrado possa realizar seus estudos mas, por outro lado, estudantes que, em muitos casos, já estão em fase adiantada de seus estudos, não são capazes de ter um desempenho considerado suficiente nesses testes de proficiência. Assim, uma interpretação simplista desses resultados põe em jogo a própria função que se espera que os testes de

* Esse trabalho foi apresentado em uma sessão de Posters Coordenados

proficiência tenham.

Portanto, parece importante investigar-se mais detalhadamente esse processo de testagem e o papel que lhe é conferido nesse contexto acadêmico. Para tanto, o presente trabalho visa abordar a questão desse acentuado índice de reprovação sob uma perspectiva que busque, tanto um maior entendimento sobre o teste de proficiência em si, quanto um exame mais atento da população envolvida. A premissa que norteia o estudo aqui reportado é de que os testes de proficiência podem efetivamente desempenhar uma função importante no contexto acadêmico da UFSM (e de outras universidades). Isso, no entanto, pressupõe que a função do teste seja considerada não somente a partir da qualidade do teste em si, mas, principalmente, de sua inserção no contexto específico dos cursos de pós-graduação.

Nesse sentido, procurar-se-á responder, ao longo do presente trabalho, se a visão dos examinandos sobre o processo de testagem a que são submetidos, importa na interpretação de seu desempenho nos testes.

1. Algumas considerações sobre testagem de proficiência

Uma das dimensões mais importantes reconhecidas na literatura sobre testagem está diretamente relacionada com seu grau de precisão como instrumento de mensuração de uma determinada habilidade lingüística (Celia, 1983: 42), ou seja, sobre sua fidedignidade. No entanto, é na demonstração quantitativa dessa dimensão do teste que se concentra parte substancial da literatura sobre fidedignidade. A preocupação com esses aspectos quantitativos tem guarida na tentativa de se garantir que um teste tenha duas características muito importantes. Primeiro, que reflita tão fielmente quanto possível as situações e contextos sobre os quais o teste se baseia. Em segundo lugar e, concomitantemente, que os escores expressos por ele sejam representativos do nível de habilidade que o teste se propõe a testar. É particularmente essa segunda característica que se discutirá aqui.

Considerando que muitas decisões acadêmicas são tomadas com base no desempenho dos candidatos nos testes de proficiência, é importante que os escores que expressam esse desempenho tenham a menor margem de erro possível, de forma que as decisões neles pautadas sejam dignas de crédito. Assim, é de fundamental importância que, no desenvolvimento e uso desses testes se tente “identificar as fontes

potenciais de erro numa dada medida da habilidade comunicativa de linguagem e minimizar o efeito desses fatores naquela medida” (Bachman, 1990: 160).

Acredita-se que basicamente quatro grupos de fatores afetem o escore de um candidato num teste: competência comunicativa, características do método de testagem, atributos pessoais (por exemplo, conhecimento prévio na área de conteúdo testada) e fatores aleatórios.

De acordo com Bachman (ibid.: 164) a influência do método de testagem no escore do indivíduo ocorreria de uma forma sistemática, já que supõe-se uma uniformidade entre diferentes edições de um teste no que concerne, por exemplo, ao tipo de questões formuladas. Convém lembrar ainda que o efeito do método de testagem sobre o escore pode ser o mesmo para todos os indivíduos testados, podendo também sofrer variações individuais. De forma similar, a influência dos atributos pessoais (não relacionados diretamente com a habilidade lingüística) sobre o escore do indivíduo também ocorreria de forma sistemática, ou seja, as características individuais, tais como estilo cognitivo, e o conhecimento prévio afetariam o desempenho do candidato no teste regularmente.

Além da competência comunicativa e desses dois grupos de fatores que ocorrem sistematicamente, os fatores do quarto grupo mencionado por Bachman afetariam o desempenho dos candidatos no teste de forma assistemática. Entre eles, pode-se citar o estado emocional do candidato e variações incontroladas nos métodos de testagem.

Em função da influência desempenhada por esses fatores, Bachman (ibid.: 226) argumenta que é “fundamental para o desenvolvimento e uso de testes de habilidade de uso de língua, ser capaz de identificar e estimar os vários fatores que atuam sobre os escores dos testes” como forma de reduzir os efeitos que não estejam diretamente relacionados com a habilidade lingüística que se esteja testando., garantindo-se assim, que os resultados expressos pelo escore do candidato no teste reflitam fidedignamente seu nível de habilidade.

Em geral, fidedignidade e validade são consideradas duas características distintas, porém relacionadas, ou seja, “embora a validade de um teste seja considerada a característica mais importante, a fidedignidade é uma condição necessária para a validade” (ibid.: 227). Em

consequência desse entendimento, Hughes (1989: 42) adverte que ao buscar maior fidedignidade nos testes desenvolvidos, devemos estar cientes do risco da redução de sua validade. Uma forma de contornar tal risco é o estabelecimento de tarefas que resultem num comportamento que demonstre que o candidato leu com sucesso (ibid.: 116).

Finalmente, é necessário observar que, em função dos diferentes fatores que podem influenciar o desempenho dos examinandos num teste, é virtualmente impossível criar um teste perfeitamente fidedigno (cf. Alderson, Clapham & Wall, 1995: 87). Assim, há que se considerar que o baixo desempenho dos examinandos no teste de proficiência possa também estar relacionado a problemas associados à fidedignidade do teste.

2. Metodologia

O presente trabalho reporta a três principais fontes de dados: 1) o teste em si, em termos do texto utilizado e das micro-habilidades de leitura (cf. Munby, 1978) aparentemente avaliadas ao longo de suas quinze questões; 2) as respostas de 61 examinandos que submeteram-se ao teste de proficiência em leitura em ESP na UFSM em abril de 1996 a um questionário entregue no recinto de aplicação do teste; e 3) o desempenho desses 61 examinandos no teste.

A correlação entre essas três fontes de dados foi feita com vistas a uma interpretação sobre o acentuado índice de reprovação observados e os principais pontos dessa análise são comentados na próxima seção.

3. Resultados e discussão

Através da análise das respostas dos 61 examinandos ao questionário foi possível averiguar, por exemplo, que os examinandos parecem reconhecer a importância da habilidade de leitura em ESP e, em função disso, da própria testagem a que são submetidos. Ao indicarem diferentes formas de comprovação da habilidade de leitura em ESP, os examinandos levam a crer que, se por um lado, a habilidade de leitura em ESP é fundamental para o desenvolvimento de seus estudos, podendo ser considerada como um pré-requisito para a aceitação de candidatos a cursos de pós-graduação (conforme indicado por 45% deles), por outro lado, as sugestões aparentemente contraditórias de que o teste de proficiência possa ser realizado durante o curso, feita por 47.5% dos

examinandos, ou de que possa ser substituído por cursos de ESP, conforme indicação de 59% dos sujeitos, podem ser interpretadas como um sinal de que eles têm consciência de que ao engajarem-se em cursos de pós-graduação vêm-se na contingência de ler em língua estrangeira e percebem que a leitura quotidiana pode melhorar seu nível de habilidade (cf. Haas, 1994), havendo, no entanto, problemas que só mesmo a instrução formal em ESP pode ajudar a resolver. O reconhecimento da potencialidade que têm para ler em ESP e das limitações que por ventura tenham, parece estar refletida na auto-avaliação que fazem de sua própria habilidade de leitura.

A maioria dos examinandos autoconceituou sua habilidade de leitura como sendo 'boa/razoável' (78.7%). Um número bem menos expressivo (11.5%) julga ter um excelente nível de habilidade de leitura e três examinandos (4.9%) julgaram que sua habilidade de leitura deveria ser conceituada como 'não satisfatória'. Além desses, outros 3 examinandos (4.9%) preferiram não opinar a respeito.

Apesar da grande maioria ter conceituado sua habilidade de leitura como sendo boa ou razoável, talvez por já estarem realizando as leituras necessárias ao desenvolvimento de seus projetos, 25 deles (41%) já haviam prestado teste de proficiência anteriormente, quase todos em inglês e com reprovação. Somado a esse fato, 28 dos 61 examinandos (45.9%) reprovaram no teste analisado.

A questão que se coloca então é: o que estaria fazendo com que examinandos reprovem num teste de proficiência cujo objetivo é aferir o nível de habilidade de leitura necessário à realização de seus cursos, se esses mesmos sujeitos já estão, de uma forma ou de outra, desenvolvendo seus projetos de pesquisa, alguns até em estágios avançados de seus cursos e que, além disso autoconceituam sua habilidade de leitura como sendo, no mínimo, boa ou razoável?

Para tentar responder a essa questão, direciono agora a exposição para o próprio teste de proficiência aplicado na UFSM em abril de 1996. Sinteticamente, o teste foi analisado a partir do texto cuja leitura era base para a solução das tarefas propostas nas questões do teste, das questões propriamente ditas e das micro-habilidades de leitura que aparentemente estavam sendo avaliadas nessas questões. A análise do teste indicou que o texto base era oriundo da área de Filosofia da Ciência e que as questões do teste aparentemente procuraram avaliar uma variedade de micro-

habilidades (cf. Munby, 1978), que incluíam leitura geral (seção **I** do teste) e seletiva (seção **III**), sumarização (seções **II** e **III**), conhecimento de elementos operantes no nível de coesão (seções **IV** e **V**) e no nível de coerência do texto (seções **I** e **III**), conhecimento lexical (seção **V**), e compreensão de significado conceitual (seção **III**). Assim, embora algumas micro-habilidades fossem foco específico de avaliação nas questões e seções do teste, freqüentemente o acesso a elas envolvia o uso de outras micro-habilidades. Isso, a meu ver, reforça a idéia de que o teste analisado se pauta numa visão de leitura como um processo complexo que envolve diferentes níveis de processamento em operação simultânea, ou seja, de que é embasado numa visão interativa de leitura (cf. Motta-Roth, 1991: 94).

A análise do teste indicou alguns problemas que podem estar relacionados à formulação das questões ou à proposição das tarefas a serem cumpridas nas questões. Nesse sentido, embora a seção **III**, por exemplo, aparentemente tenha tentado avaliar a habilidade dos examinandos de localizar pontos específicos no texto, além da habilidade de distinguir entre a idéia principal e os detalhes secundários, não se pode ser incisivo em afirmar que o baixo desempenho dos examinandos esteja relacionado apenas à falta dessas duas habilidades. Talvez haja mais fatores envolvidos, como, por exemplo, as restrições na interação do leitor com o texto, que podem ter causado falhas na atribuição de relevância às partes do texto pelo leitor (Cavalcanti, 1989: 46-47) ou até mesmo o nível de dificuldade das questões.

Um outro problema que pode ser apontado em relação ao teste diz respeito aos níveis extremos de acertos em algumas questões. As questões das seções **I** e **II**, por exemplo, têm um nível muito elevado de acertos para todos os examinandos, enquanto que nas questões das seções **III** e **V** esse nível é consideravelmente baixo. A diferença entre esses dois grupos de questões é que naquelas em que o nível de dificuldade foi maior, ainda é possível fazer alguma distinção entre o desempenho de examinandos com diferentes autoconceitos sobre sua habilidade de leitura, ao passo em que nas primeiras isso praticamente não ocorre. Seguindo a mesma linha de raciocínio, pode-se dizer que as questões da seção **IV** são aquelas que permitem diferenciar melhor o desempenho de examinandos com diferentes avaliações sobre sua habilidade de leitura.

A conseqüência dessa alta concentração tanto de acertos, como de erros é que tais questões podem estar informando muito mais sobre o

nível de facilidade/dificuldade das questões do que propriamente sobre a habilidade de leitura dos examinandos, além de dificultar a percepção sobre a influência do nível de especificidade de conteúdo do teste sobre o desempenho dos examinandos.

4. Observações finais

Procurando, portanto, responder à questão proposta no título deste trabalho, tentou-se mostrar que, embora a análise do teste em si seja importante no que concerne aos parâmetros técnicos envolvidos na testagem, a consideração de informações provenientes dos examinandos pode contribuir para a ampliação do entendimento que se tem sobre o teste e, principalmente, sobre o contexto onde essa testagem se desenvolve. Pode contribuir, por exemplo, na avaliação da credibilidade dos processos de testagem desenvolvidos em nossas universidades (referida na literatura como validade aparente). Apesar da alegação de falta de cientificidade desse tipo de validade (ver Wielewicki, 1997a), um teste que não seja considerado por seus usuários como uma medida representativa da habilidade que se procura medir pode ter problemas relacionados à motivação e ao empenho dos examinandos no teste.

Finalmente, a consideração da visão dos examinandos sobre leitura e testagem também pode ser de valia para uma melhor compreensão da relação recíproca e contínua entre ensino e testagem (ver Alderson & Wall, 1993: 115) e sobre a própria questão de qual o nível de especificidade de conteúdo que um teste deva ter.

REFERÊNCIAS BIBLIOGRÁFICAS:

- ALDERSON, J.C. & D. WALL. (1993). Does washback exist? *Applied Linguistics*. Vol. 14, n. 2, pp. 115-29.
- ALDERSON, J.C., C. CLAPHAM & D. WALL. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- BACHMAN, L.F. (1990) *Fundamental considerations in Language Testing*. Oxford: Oxford University Press.
- CAVALCANTI, M. C. (1989) *Interação leitor texto: Aspectos de interpretação pragmática*. Campinas: Editora da UNICAMP.
- CELIA, M.H.C. (1983) *Leitura em inglês: formato alternativo para um exame de proficiência*. Porto Alegre: Curso de Pós-Graduação em Lingüística e Letras da Pontifícia Universidade Católica do Rio

- Grande do Sul. Dissertação de Mestrado.
CONSELHO FEDERAL DE EDUCAÇÃO. Parecer n.º 977/65 de 03 dez. 1965. Relator: Newton Sucupira. Normas para regulamentação dos cursos de pós-graduação. *Documenta*, Rio de Janeiro (44): 67-86. 20 p.
- HAAS, C. (1994) Learning to read biology: one student's rhetorical development in college. *Written Communication*. vol. 11, n.1. p. 43-84.
- HUGHES, A. (1989) *Testing for language teachers*. Cambridge: Cambridge University Press.
- MOTTA-ROTH, D. (1991) O processamento de sentido na leitura de textos em inglês como língua estrangeira. *Letras*. vol. 1. Santa Maria: UFSM/CAL. p. 92-102.
- MUNBY, J. (1978) *Communicative syllabus design: a sociolinguistic model for defining the content of purpose-specific language programs*. Cambridge: Cambridge University Press.
- WIELEWICKI, H.G. (1997a) *Face validity in EAP reading proficiency tests: Does it matter?* Trabalho apresentado na 1997 Annual Conference of the American Association for Applied Linguistics. March 8-11, Florida, USA.
- (1997b). *Testagem de proficiência em leitura em inglês: examinandos e teste como fontes de entendimento sobre esse processo*. Santa Maria, RS: Curso de Mestrado em Letras, Universidade Federal de Santa Maria. Dissertação de Mestrado.