

## Resumo

### Assessment of Reading in English as a Foreign Language: Investigating the Defensibility of Test Items

Celso Henrique Soufen Tumolo

Universidade Federal de Santa Catarina  
2005

Profª. Dra. Lêda Maria Braga Tomitch  
Orientadora

Neste estudo, a defensabilidade de itens de testes é investigada em três situações de teste: testes de proficiência, testes de sala de aula, e vestibulares. A defensabilidade é julgada recorrendo-se ao conceito de validade proposto por Messick (1989) para testes em educação e por Bachman (1990) para testes em línguas, isto é, em termos da validade das interpretações e ações baseadas nas respostas aos itens de testes, considerando as justificativas oriundas das análises de evidências e de conseqüências. Para a investigação de evidências, para evidências relativas ao construto, construtos de habilidade lingüística e de habilidade de leitura são usados, e para evidências relativas ao critério, o critério definido para estudos acadêmicos proposto por Weir, Huizhong, e Yan (2000) e o arcabouço proposto por Bachman (1990) para a avaliação da autenticidade de tarefas presentes nos testes em relação às tarefas exigidas para estudos acadêmicos são usados. Para a investigação das conseqüências, uma apreciação das conseqüências é calcada na noção de justo de Bachman (1990) e na perspectiva crítica do uso de testes de Shohamy (2001). O método usado para a investigação é pautado na noção recente de validade baseada em argumentos, proposta por vários pesquisadores na área de avaliação. É também pautado na tabela de validade proposta por Chapelle (1994) para considerações das forças dos argumentos, tanto a favor como contra a interpretação de habilidade baseada nos itens, com uma terceira coluna adicional com a refutação do argumento contra, como sugerido por Davidson (pessoalmente, 2004). Considerando os argumentos para cada item, é possível concluir que há itens defensáveis e itens não defensáveis. Alguns dos itens defensáveis focam em habilidade como identificação de elementos gramaticais e

de coesão, inferência de significado de palavras desconhecidas, inferências elementares e proposicionais, e identificação do valor funcional do uso da língua. Alguns itens não-defensáveis têm problemas técnicos, como também aspectos irrelevantes ao construto de leitura de avaliar outros construtos, como conhecimento de vocabulário, conhecimento de mundo, habilidade de escrita, compreensão de ilustrações, como também aspectos irrelevantes ao critério de avaliar tarefas não exigidas para estudos acadêmicos, como tarefas específicas para leitura de textos não-acadêmicos como poemas, narrativas, propaganda comercial, todos dificultando que o desempenho nos testes reflita a habilidade avaliada e auxilie na avaliação de desempenho futuro. Outros itens não-defensáveis têm conseqüências de favorecer alguns grupos em detrimento de outros, já que conhecimento específico é pressuposto para algumas de suas respostas. As implicações pedagógicas e políticas das conclusões são que a escolha e desenvolvimento de itens de testes incorporem noções de validade, em todas as suas facetas, considerando tanto a validação de construto, como também autenticidade das tarefas e impacto do uso do teste na sociedade e nos indivíduos, de tal forma a tornar os itens usados mais defensáveis.