



<http://dx.doi.org/10.30681/real.v11i2.2720>

ABRANGÊNCIA DO TERMO “ESTRATÉGIAS” NO EFEITO RETROATIVO EM EXAME DE PROFICIÊNCIA EM LÍNGUA INGLESA

Eliana KOBAYASHI (IFSP)¹

Resumo: O efeito retroativo que testes de proficiência em língua inglesa de alta relevância pode causar no ensino tem sido muito discutido nos estudos sobre avaliação e mostram possibilidades de consequências positivas e também negativas. Muitas vezes, o termo estratégias ou técnicas para fazer um teste é visto de forma negativa pelos atores envolvidos, como se fosse um aspecto que não tivesse relação com as habilidades de uso da língua. O objetivo principal deste artigo é focalizar dois termos pouco explorados quando o efeito retroativo é analisado: a experiência em responder teste (*test-wiseness*) e estratégias para a realização de testes (*test-taking strategies*) demonstrando que, muitas vezes, o limite entre eles não está claramente estabelecido. Além disso, busca-se também ressaltar aspectos positivos presentes em ambos pelo aprofundamento de suas definições.

Palavras-chave: Efeito retroativo. Experiência em responder testes. Estratégias para a realização de testes.

Abstract: The washback effect of high-stakes EFL proficiency tests in teaching has been much discussed in the assessment field showing possibilities of positive and also negative consequences. The terms strategies or techniques are many times seen as an aspect not related to the language use development. The main objective of this paper is to focus on two terms not so discussed when analyzing the washback effect: the test-wiseness and test-taking strategies showing that the boundary between them is many times blurred. In addition, the paper attempts to present their positive aspects by deepening their definitions.

Keywords: Washback effect. Test-wiseness strategies. Test-taking strategies.

1. Introdução

Testes são utilizados na sociedade com diversos objetivos e em diferentes contextos. De provas escolares a seleção de empregos, esses instrumentos de avaliação surgem em vários momentos da vida de uma pessoa. Entretanto, a relevância que apresentam pode variar. De acordo com Madaus (1988), caso os resultados do teste sejam utilizados para a tomada de decisões que afetem a vida dos candidatos, são considerados de alta relevância. Assim, dependem das percepções dos envolvidos, que incluem também professores e instituições de ensino. Por exemplo, se a obtenção de um emprego depende da aprovação em um teste de proficiência em inglês, como o *Test of English for International Communication* (TOEIC), muito utilizado no meio empresarial, o insucesso

¹ Eliana Kobayashi: Doutora em Linguística Aplicada (Unicamp). Instituto Federal São Paulo. likobayashi@yahoo.com



no exame² pode afetar negativamente a vida do candidato. De forma semelhante, mas no meio acadêmico, muitos alunos que buscam participar de intercâmbios em universidades internacionais, também devem ser aprovados em testes de inglês, como o *Test of English as a Foreign Language* (TOEFL).

Portanto, é possível verificar que devido à importância que testes de proficiência em língua inglesa podem representar na vida das pessoas, torna-se esperado que tais exames causem efeito retroativo no ensino e aprendizado.

2. Efeito retroativo

O efeito retroativo que um exame de proficiência em língua inglesa pode causar no processo de ensino e aprendizagem, no sistema educacional, nos atores envolvidos e na sociedade tem sido estudado há muito tempo. Assim, muitas definições buscam abranger os vários aspectos desse fenômeno linguístico. Buck (1988) afirmava que efeito retroativo era a influência do teste nas atividades que alunos e professores realizavam, principalmente se o exame fosse importante e o índice de aprovação medisse o sucesso do docente. Outros estudos também apontavam a existência do efeito retroativo em sala de aula como de Paris, Lawton, Turner e Roth (1991), Haladyna, Nolen e Haas (1991), Frederiksen (1984) e Vernon (1956).

Por outro lado, Alderson e Wall (1993) apontaram para a necessidade de evidências empíricas dos mecanismos do efeito, o que as pesquisas até então careciam. Similarmente, Messick (1994) trouxe contribuições importantes para a concepção do fenômeno, afirmando que efeito retroativo refere-se à extensão em que um exame influencia os professores e alunos a fazerem algo que não fariam necessariamente sem o exame. Assim, um efeito só poderia ser considerado retroativo se houvesse evidências relacionadas à introdução e uso do teste.

O efeito retroativo tem o potencial de ser positivo ou negativo, sendo essa dimensão intitulada de valor por Watanabe (1997). Messick (1996) afirmava que para se promover o efeito positivo, as atividades desenvolvidas para a aprendizagem da língua e aquelas destinadas a preparação para o teste não deveriam ser distintas. Similarmente, Pearson (1988) argumentava que testes bons podem ser diretamente utilizados como atividades de ensino e aprendizagem. Além disso, Davies (1985) defendia que um exame

² Neste artigo, os termos exame e teste são utilizados como sinônimos.



criativo e inovador tem o potencial de provocar mudanças no currículo ou em um novo currículo, assim como influenciar os testes de rendimento.

Por outro lado, muitos aspectos negativos do efeito retroativo também já foram apontados como: estreitamento de currículo, ensino voltado para o teste, foco no ensino das disciplinas avaliadas pelo exame em detrimento de outras, uso inadequado do tempo em sala para favorecer técnicas para o exame, pressões sobre os professores para aprovação, problemas de falta de validade e injustiça do teste. (VERNON, 1956; WISEMAN, 1961; POPHAM, 1987; SMITH, 1991; ALDERSON, WALL, 1993; HAYES, READ, 2004; STECHER, CHUN, BARRON, 2004; FERMAN, 2004).

Para Alderson e Wall (1993), efeito negativo consistia no efeito indesejável no ensino e aprendizado de exame considerado ruim, ou seja, algo que o professor ou o aluno não deseja ensinar ou aprender. Noble e Smith (1994) identificaram como efeitos negativos no ensino de testes de alta relevância: técnicas para a realização de testes e práticas repetitivas de questões de múltipla escolha visando uma pontuação mais alta que, por outro lado, dificilmente promovem compreensão.

Quando focalizamos a preparação para um teste de proficiência em língua inglesa considerado de alta relevância, é muito comum encontrarmos o termo técnica ou estratégia, como citado por Noble e Smith (1994), frequentemente com uma conotação negativa, distante de uma visão de ensino de língua. Por outro lado, é importante questionarmos, primeiramente, a sua definição e se realmente esse termo pode ser sempre considerado em toda a sua abrangência como um dos aspectos do efeito retroativo negativo de um teste.

No Brasil e no mundo, os cursos preparatórios para testes de proficiência são muito comuns e uma das visões mais marcantes comumente associadas consiste no ensino de estratégias para a realização do exame em questão, levantando a possibilidade de que aqueles alunos que não têm interesse em realizar tal teste não poderiam ser beneficiados em seu desenvolvimento linguístico frequentando esse curso. Notamos também que para alguns professores que ministram esses cursos, o termo estratégia pode surgir dissociado das ações que levariam ao ensino da língua, como mostram os resultados do estudo de Kobayashi (2016), indicando a possibilidade de efeito retroativo negativo do exame. No entanto, acreditamos que há uma carência tanto da discussão sobre as definições dessas estratégias mencionadas tão frequentemente quanto da análise sobre a relação entre essas estratégias e o desenvolvimento das habilidades de uso da língua. Primeiramente, é necessário diferenciar os vários aspectos que integram o termo estratégia e, após isso,



buscar estabelecer a relação mantida com o ensino da língua.

3. Estratégias relacionadas à experiência em responder testes (*test-wiseness strategies*)

Podemos considerar como esperado que, em contextos de ensino voltados a exames, haja a utilização de todos os recursos disponíveis e esforço dos atores, como direção da instituição de ensino, coordenadores, professores, alunos e de suas famílias, entre outros para que os candidatos sejam aprovados nos exames. Tais recursos podem estar ligados diretamente ao conteúdo avaliado pelo exame ou ser externos a ele.

Um termo frequentemente citado nesses contextos é o *test-wiseness strategies* que, segundo tradução de Almeida Filho e Schmitz (1998), trata-se das estratégias relacionadas à experiência em responder testes. A discussão sobre a sua definição, como uma variável passível de influenciar os resultados em um teste, foi iniciada por Thorndike (1951).

Astúcia para saber quando adivinhar e um olho preciso para dicas secundárias e dicas externas ao conteúdo do teste que são possíveis de serem usadas em uma grande variedade de testes, especialmente aqueles que não estão bem construídos (...) geralmente representa uma variação inválida sistemática funcionando para reduzir sistematicamente a validade do teste. (THORNDIKE, 1951, p. 569)

Desde então, muitas outras definições similares surgiram (GIBB, 1964; DIAMOND, EVANS, 1972; DOLLY, WILLIAMS, 1983) sendo bastante citada a de Millman (1965), para quem experiência em se fazer teste refere-se à capacidade de uma pessoa em utilizar as características e formatos de um teste ou/e das situações de realização do teste para obter notas mais elevadas, sendo isso logicamente independente do assunto que as questões supostamente avaliam. Na mesma linha, segundo Benson (1988), trata-se da habilidade cognitiva ou um conjunto de estratégias que o candidato pode utilizar para melhorar os resultados em um teste independente da sua área de conhecimento. Cohen (1995) também afirma que as estratégias da experiência em responder testes não são necessariamente determinadas pela proficiência na língua que está sendo avaliada, mas pode ser dependente do conhecimento do candidato sobre como realizar testes.

Diversas taxonomias foram propostas para essa experiência em responder testes. Sarnaki (1979) apresentou uma taxonomia de cinco estratégias: 1) estratégias de utilização do tempo: começar a realizar o teste rapidamente com certa precisão; 2) estratégias para se



evitar erros: por exemplo, prestar atenção nas instruções; 3) estratégias de adivinhação: adivinhar somente quando não há penalização para tanto; 4) estratégias de raciocínio dedutivo: fazer uso das informações de conteúdo do teste para outras questões e opções; 5) estratégias de dicas e intenções: reconhecer e fazer uso de qualquer idiossincrasia relevante do teste que diferencia a resposta correta da incorreta.

Nitko (2001), por outro lado, reduziu a taxonomia de Sarnaki (1979) e propôs: 1) estratégias de utilização do tempo; 2) estratégias para se evitar erros; e 3) estratégias de adivinhação.

Além disso, Watter e Siebert (1990) dividem as estratégias de experiência em responder testes em três categorias principais: 1) estratégias utilizadas antes de iniciar o teste: ler todas as questões e começar com as mais fáceis; alocar tempo para cada questão conforme dificuldade e extensão; ler as instruções cuidadosamente; sublinhar palavras-chaves. 2) estratégias usadas durante o exame: responder as questões em ordem cronológica, revisar as questões logo depois de respondê-las; utilizar todo o tempo disponível; responder imediatamente o que vier à cabeça; responder todas as questões até as que não sabe. 3) estratégias usadas depois de terminar o teste: revisar as respostas para corrigir erros gramaticais e de ortografia, reler as questões para verificar se as instruções foram entendidas corretamente; evitar correções de última hora.

Assim, como Stanley (1971) e Crehan *et al* (1974) afirmam, a experiência em responder testes está relacionada a capacidade, traços, atributos ou estados individuais do que caracteriza os exames. Entretanto, Diamond e Evans (1972) alertam que essa experiência não deveria ser vista somente como uma característica do candidato, mas sim como uma idiossincrasia psicométrica dos testes relacionados a dicas em itens que apresentam falhas. Dessa maneira, um item de teste que apresenta falhas na sua elaboração gera possibilidades para que um candidato utilize habilidades e conhecimentos não associados ao conteúdo para resolver a questão.

Sintetizando esses posicionamentos, Millman *et al* (1965) afirmaram que a experiência em responder um teste pode ser dividida em duas partes: 1) estratégias que dependem do candidato; e 2) estratégias dependentes da elaboração do teste e/ou propósitos do teste. Essa é a visão que tem sido amplamente aceita (ROGERS, YANG, 1996).

Por outro lado, há controvérsias na área, como o questionamento de Rogers e Bateson (1991) sobre a dependência ou independência da experiência em responder testes do assunto que está sendo avaliado.

(...) parece que a aplicação efetiva das estratégias da experiência em fazer teste



é dependente de um conhecimento parcial (do assunto avaliado no teste). Esse conhecimento parcial, embora insuficiente para responder a questão do teste, é suficiente quando acoplado com o conhecimento dos princípios da experiência em fazer teste para aumentar a probabilidade de responder corretamente os itens suscetíveis a tais princípios. Alunos com baixo conhecimento do conteúdo, mas com experiência em fazer teste e alunos com alto conhecimento do conteúdo, mas baixa experiência em fazer teste terão desempenho menor em tais itens do que os alunos que possuem ambos. (ROGERS, BATESON, 1991, p. 210)

Podemos verificar que há muitas discussões sobre a dependência ou não do conteúdo, embora mais comumente a experiência em responder teste não esteja relacionada ao conteúdo do teste. No entanto, acreditamos que o mais relevante nessa discussão é a complexidade em se estabelecer limites de definição e também levantar o questionamento se o desenvolvimento de todos os aspectos que abrangem essa experiência pode ser considerado como efeito retroativo negativo de um teste, ou seja, se, ao invés de desenvolver as habilidades de uso da língua alvo, professores, alunos e demais envolvidos poderiam focalizar estratégias que não estabelecem relação alguma com o saber utilizar uma língua somente visando notas mais elevadas no exame.

Um exemplo dessa situação seria as estratégias de pós-teste apontadas por Watter e Siebert (1990) como revisar as respostas para corrigir erros gramaticais e de ortografia, reler as questões para verificar se as instruções foram entendidas corretamente. Acreditamos que, na medida em que produzir um texto escrito na vida real exige revisões de ortografia e gramática, seria importante um aluno aprender tais estratégias para o desenvolvimento de suas habilidades linguísticas, principalmente de compreensão e produção escrita. Dessa forma, tal estratégia não poderia ser considerada um aspecto negativo do efeito retroativo de um teste. Outro exemplo similar seria ler as instruções de uma questão e sublinhar as palavras chaves, consideradas estratégias pré-teste, pelos autores. No entanto, não acreditamos que tais estratégias sejam distintas daquelas comumente utilizadas em sala de aula, uma vez que sublinhar palavras chaves em um texto, seja uma instrução ou qualquer outro gênero, para facilitar a sua compreensão consiste em um dos passos da técnica de leitura.

Por outro lado, se um teste de proficiência tenha sido elaborado de forma que permita os chamados macetes, concordamos com Diamond e Evans (1972), pois há um problema na sua elaboração que causa um efeito negativo no ensino, visto que alunos começam a treinar para o teste sendo que deveriam desenvolver suas habilidades.

Portanto, podemos verificar que a própria abrangência do termo da experiência



em responder teste não permite afirmações categóricas sobre a natureza do seu efeito no ensino.

4. Estratégias para a realização de testes (*test-taking strategies*)

Quando citadas por atores, as estratégias para a realização de testes muitas vezes assemelham-se às estratégias de experiência em responder testes, outras vezes aparentam estar relacionadas a maneiras de lidar com algum tipo de questão. Entretanto, consideramos que essa distinção deve ser feita, para melhor compreensão de sua abrangência dentro do conceito de efeito retroativo.

Baseamos muito da nossa discussão no trabalho de Cohen (1998), que estabelece que estratégias de uso da língua e estratégias para a realização de testes consistem nos mesmos processos:

Estratégias de uso da língua são operações mentais ou processos que os aprendizes conscientemente selecionam quando estão executando uma tarefa de língua. Essas estratégias também constituem as estratégias para a realização de testes quando elas são aplicadas a tarefas em testes de língua (...) nessa discussão serão vistas como aqueles processos que os respondentes selecionaram de forma consciente (...) a noção de estratégia implica um elemento de seleção, caso contrário, esses processos não seriam considerados estratégias. (COHEN, 1998, p.92)

O autor ressalta que não se deve assumir que alguma estratégia seja boa ou ruim para uma determinada tarefa, pois isso depende de como o candidato, que possui um perfil cognitivo específico e seu próprio repertório de estratégias para fazer o teste, emprega as estratégias diante de uma determinada tarefa e em um dado momento. Por isso, há candidatos que conseguem ter sucesso nos testes com um número limitado de estratégias, fazendo um bom uso delas na maior parte, enquanto outros podem saber muitas, mas não utilizá-las de modo eficiente.

Conforme Cohen (1998, p.93), a capacidade de usar as estratégias de língua tem sido referida como competência estratégica, um componente do uso comunicativo de língua, integrante do modelo teórico proposto por Canale e Swain (1980): “este modelo enfatiza as estratégias compensatórias, isto é, estratégias utilizadas para compensar ou remediar carências em alguma área da língua”.

O autor acrescenta que Bachman e Palmer (1996) ampliaram esse modelo,



passando a incluir três componentes: definição de objetivos, avaliação e planejamento. Para os autores, competência estratégica constitui em um conjunto de componentes metacognitivos ou estratégias consideradas processos executivos de ordem elevada que funcionam como gerenciamento cognitivo no uso da língua, assim como em outras atividades cognitivas. Assim, usar uma língua envolve o conhecimento do assunto, esquemas afetivos e conhecimento da língua, que permitem que o usuário crie e interprete discursos de modo adequado à situação.

Além disso, Bachman e Palmer (1996) enfatizam que, em avaliação de línguas, a competência estratégica proporciona uma base fundamental para o desenvolvimento de tarefas interativas para o teste e para avaliar a interatividade das tarefas que são usadas, e identificam três áreas nas quais os componentes de estratégias metacognitivas operam:

1. Definição de objetivos: decidir o que o usuário vai fazer. Identificar as tarefas do teste; escolher uma ou mais tarefas de um conjunto de tarefas, se for dada essa escolha; decidir se vai ou não tentar completar a tarefa. Para Bachman e Palmer (1996), como o propósito do exame de línguas é extrair uma amostra específica de uso da língua, há apenas uma variedade limitada de tarefas, assim a flexibilidade do candidato na definição de objetivos para realizar as tarefas não é tão grande quanto em situações de não teste.
2. Avaliação: analisar o que é preciso, com o que o candidato deve trabalhar e o quão bem realizou. Os autores estabelecem três divisões para essa estratégia. 1) Determinar o desejo e viabilidade de completar as tarefas com sucesso e quais elementos do conhecimento da língua e do assunto são necessários; 2) Avaliar o seu próprio conhecimento do assunto e de língua para completar a tarefa com sucesso; 3) Avaliar a adequação da resposta à tarefa do teste. O critério relevante dessa estratégia está relacionado às características gramatical, textual, funcional e sociolinguística da resposta, assim como o conteúdo.
3. Planejamento: decidir o que fazer com o que o candidato tem. Essa estratégia envolve três aspectos: 1) selecionar um conjunto de elementos específicos do conhecimento do tópico e de língua que será usado no plano; 2) formular um ou mais planos para a resposta; 3) selecionar um plano de implementação como resposta à tarefa.

A taxonomia de estratégia de realização de teste proposta por Cohen (1998), que a define fazendo referência às estratégias do aprendiz de segunda língua, está dividida em duas categorias principais: as estratégias de aprendizado de segunda língua e estratégias de uso de segunda língua, sendo que ambas consistem em ações selecionadas pelos aprendizes de forma consciente para melhorar tanto o aprendizado quanto o uso da língua.



Especificamente sobre as estratégias de uso da língua, o autor aponta quatro tipos: estratégias de resgate, usada para trazer o material linguístico armazenado; estratégias de treinamento, utilizadas para treinar as estruturas alvo da língua; estratégias de proteção, usadas para criar a impressão, não verdadeira, de que possuem controle sobre o material linguístico; e as estratégias de comunicação, para transmitir uma mensagem significativa e informativa.

Essas quatro estratégias, de acordo com Cohen (1998), seriam as mesmas praticadas durante a realização de um teste visto que o candidato necessita resgatar o material linguístico para a sua utilização e também possivelmente praticá-los.

As taxonomias de Cohen e o modelo de Bachman e Palmer demonstram que as estratégias para a realização de testes não estão dissociadas do uso da linguagem; pelo contrário, estão extremamente relacionadas e, portanto, caso o construto do teste reflita o uso do inglês em situações próximas do real, as estratégias para a realização do teste, as estratégias de uso da língua e realização do teste estariam ainda mais relacionadas.

Dessa forma, ressaltamos que ao se ouvir o termo estratégia relacionada a cursos ou aulas de preparação para exames de proficiência, é preciso primeiramente identificar a qual definição o termo está mais relacionado: à experiência em responder testes (*test-wiseness strategies*) ou à estratégia para a realização de testes. Em seguida, analisar também se o aspecto identificado está realmente desvinculado de uma ação de desenvolvimento da língua alvo. Somente assim, será possível avaliar essa dimensão de valor do efeito retroativo do teste em questão.

5. Considerações finais

Este artigo trouxe a discussão de dois termos relacionados à dimensão de valor do efeito retroativo de um teste de proficiência em língua inglesa de alta relevância: a experiência em responder testes (*test-wiseness strategies*) e a estratégia para a realização de testes (*test-taking strategies*) em uma tentativa de diferenciar e aprofundar o termo “estratégias para o teste” comumente utilizado na preparação e realização de qualquer exame. Para tanto, estruturamos nosso argumento na relação existente entre: testes de proficiência em língua inglesa de alta relevância, efeito retroativo e sua dimensão de valor, experiência em responder teste e estratégias para a realização de testes.

Nosso objetivo foi demonstrar a relação direta que as estratégias para a realização



de testes têm com o desenvolvimento do uso da língua e que, portanto, não devem ser considerados como um aspecto negativo do efeito retroativo. Além disso, também questionamos que algumas características da experiência em responder teste também podem estar ligadas ao desenvolvimento de habilidades, apresentado, desse modo, limites não muito claros com a definição de estratégias para a realização de testes.

Referências

- ALDERSON, J. C., WALL, D. Does washback exist?, **Applied Linguistics**, 14, p.115, 1993.
- ALMEIDA FILHO, J. C.P., SCHMITZ, J. R. **Glossário de Linguística Aplicada**. Campinas: Pontes, 1998.
- BACHMAN, L., PALMER, A. **Language testing in practice: Designing and developing useful language tests**. Oxford: OUP, 1996.
- BENSON, J. The psychometric and cognitive aspects of test-wiseness: a review of the literature. In M. KEAN (ed.) **Test-wiseness**. Bloomington: Phi Delta Kaplan, 1988.
- BUCK, G. Testing Listening Comprehension in Japanese University Entrance Exams. **JALT Journal**, v.10, p.15-42, 1988.
- CANALE, M., SWAIN, M. Theoretical bases of communicative approaches to second language teaching and testing. **Applied Linguistics**, v.1, n.1, 1980.
- COHEN, A. **Strategies in learning and using a second language**. New York: Addison Wesley Longman Inc, 1998.
- COHEN, A.D. The coming of Age of Research on Test-Taking Strategies. **Language Assessment Quarterly**, v.3(4), p.307-331, 2006.
- CREHAN, K., KOEHIER, R.A., SLAKTER, M. J. Longitudinal studies of test-wiseness. **Journal of Educational Measurement**. v.11, p.209-212, 1974.
- DAVIES, A. Follow my leader: Is that what language tests do. In: LEE, Y.P., FOK, C.Y.Y., LORD, H., LOW, G. (eds). **New directions in language testing**. Oxford: Pergamon Press, 1985.
- DIAMOND, J. J., EVANS, W. J. An investigation of the cognitive correlates of test-wiseness. **Journal of Educational Measurement**, v. 9, p.145-150, 1972.
- DIAMOND, J. J., EVANS, W. J. An investigation of the cognitive correlates of test-wiseness. **Journal of Educational Measurement**, v. 9, p.145-150, 1972.



DOLLY, J.P., WILLIAMS, K.S. **Teaching test wiseness**. Paper presented at the annual meeting of the Rocky Mountain Educational Research Association. Jackson, Wyoming, 1983.

FERMAN, I. The Washback of an EFL National Oral Matriculation Test to Teaching and Learning. In: CHENG, L., WATANABE, Y., CURTIS, A. **Washback in language testing: Research contexts and methods**. Mahwah, NJ: Lawrence Erlbaum, 2004.

FREDERIKSEN, N. The real test bias. **American Psychologist**. March, 193-202, 1984.

GIBB, B. C. **Test-wiseness as a secondary cue response**. Unpublished doctoral dissertation. Stanford University, California, 1964.

HALADYNA, T.H., NOLEN, S.B., HAAS, N.S. Raising standardized achievement test scores and the origins of test score pollution. **Educational Researcher**, v.20, p-2-7, 1991.

HAYES, B., READ, J. Test preparation in New Zealand: preparing students for the IELTS academic module. In CHENG, L., WATANABE, Y., CURTIS, A. (eds). **Washback in Language Testing Research Contexts and Methods**. Lawrence Erlbaum Associates, New Jersey, p. 97-111, 2004.

KOBAYASHI, E. **Efeito retroativo de um exame de proficiência em língua inglesa em um Núcleo de Línguas do Programa Inglês sem Fronteiras**. Tese (Doutorado em Linguística Aplicada) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2016.

MADAUS, G. Public policy and testing profession. **Educational measurement: issues and practice**, v.4, p.5-11, 1985.

MESSICK, S. Validity and washback in language testing. **Language testing**, v.13, p.241-256, 1996.

MILLMAN, J., BISHOP, H., EBEL, R. An analysis of test-wiseness. **Educational and Psychological Measurement**, v. 25, p.707-726, 1965.

MILLMAN, J., BISHOP, H., EBEL, R. An analysis of test-wiseness. **Educational and Psychological Measurement**, v. 25, p.707-726, 1965.

NITKO, J. **Educational assessment of students**. New York: Merrill Prentice Hall, 2001.

PARIS, S.G., LAWTON, T. A., TURNER, J. C., ROTH, J. L. A developmental perspectives on standardized achievement testing. **Educational Researcher**, v. 20, p. 12-19, 1991.

PEARSON, I. Tests as Levers for Change. In D. CHAMBERLAIN, R.J. BAUMGARDNER (eds). **ESP in the Classroom: Practice and Evaluation**. **ELT**



- Documents**, v.28, p. 98-107. London: Modern English Publications, 1988.
- POPHAM, W.J. The merits of measurement-driven instruction, **Phi Delta Kappa**, v.68, p. 679-682, 1987.
- ROGERS, W., BATESON, D. Verification of a model of test-taking behavior of high school seniors. **Journal of Experimental Education**, v. 59: 331-349, 1991.
- ROGERS, W. T., YANG, P. Testwiseness: its nature and application. **European Journal of Psychological Assessment**. Tiburg, Netherland, 1996.
- SARNACKI, R. An examination of test-wiseness in the cognitive domain. **Review of Educational Research**, v.49: 252-279, 1979.
- SMITH, M. L. Meanings of test preparation. **American Educational Research Journal**, v.28, p.21-42, 1991.
- STANLEY, J. C. **The convergence strategy o f test-wiseness**. Paper presented at the annual meeting of the American Educational Research Association, Boston, 1971.
- STECHER, B., CHUN, T., BARRON, S. (2004). The effects of assessment-driven reform on the teaching of writing in Washington State. In: L. CHENG, Y. WATANABE, A. CURTIS (eds.). **Washback in language testing: Research contexts and methods**, p. 53-71. Mahwah, NJ: Lawrence Erlbaum, 2004.
- THORNDIKE, E. L. Reliability. In: LINDQUIST, E.F. (ed.). **Educational Measurement**, Washington, DC: American Council on Education, 1951.
- VERNON, P. E. **The Measurement of Abilities** (2nd edition). London: University of London Press, 1956.
- WATTER, T., SIEBERT, A. **Students' success: how to succeed in college and still have time for your friends**. New York: Holt, Rinehart and Winston, Inc., 1990.
- WATANABE, Y. **Washback effects of the Japanese university entrance examination: classroom- based research**. Unpublished doctoral dissertation, Lancaster University, UK, 1997.
- WISEMAN, S. (ed.). **Examinations and English education**. Manchester, England: Manchester University Press, 1961.