

Funcionamento diferencial dos itens do Teste Cloze por Opção

Maria Cristina Rodrigues Azevedo Joly

Docente. Programa de Pós-Graduação em Psicologia. Universidade São Francisco (USF). Itatiba. São Paulo. Brasil.

mcrisjoly@gmail.com

Nayane Martoni Piovezan

Doutoranda. Programa de Pós-Graduação em Psicologia. Universidade São Francisco (USF). Itatiba. São Paulo. Brasil.

nanny_mp@yahoo.com.br

Funcionamento diferencial dos itens do Teste Cloze por Opção

Resumo

O estudo visou analisar o funcionamento diferencial dos itens do Teste de Cloze por Opção EM/U o qual avalia compreensão em leitura. Participaram 489 estudantes. Destes, 206 universitários de instituição particular, sendo 79,1% do gênero feminino, com idade média de 25,89 anos ($DP = 6,79$). Outros 283 alunos do Ensino Médio (EM), com média de idade de 15,98 anos ($DP = 1,27$), 52,7% de escola pública e 54,8% do gênero feminino. O teste foi aplicado de forma coletiva. Constatou-se que quatro itens (10%) tiveram os maiores índices de dificuldade. Verificou-se apenas um item (2,5%) fora do limite adequado para *outfit*. Constatou-se que o conjunto de itens desse teste está ajustado do ponto de vista estrutural e que é relativamente fácil para esses estudantes respondê-lo. Doze itens (30%) estão funcionando diferencialmente para o EM e Superior. Novos estudos são necessários para ampliar o conhecimento acerca do teste e adequação à população a que se destina.

Palavras-chave: Funcionamento Diferencial do Item; Compreensão de leitura; Teoria de Resposta ao Item.

Differential Item Functioning of Teste Cloze por Opção

Abstract

This study analyzed the differential item functioning of Teste Cloze por Opção EM/U that assess collectively reading comprehension. The subjects were 489 students. 206 of them studied in a private college and 79,1% of these participants were female with average of 25,89 years ($SD = 6,79$). The other students were 283 from public (52,7%) and private (47,3%) high school with 15,98 years ($SD = 1,27$) as average and 54,8% of them were female. The results showed that four items (10%) had the highest rates of difficulty. There was only one item (2,5%) out of range suitable outfit. It was found that the set of items of this test is adjusted for the structural theoretical model and it is relatively easy for these students respond to it. Twelve items (30%) are functioning differentially for high school and college. Further studies are needed to enhance the test properties and his adequacy for the intended population.

Keywords: Differential Item Functioning; Reading comprehension; Item Response Theory.

Funcionamiento diferencial del ítem y análisis del modelo de Rasch de Teste Cloze por Opção

Resumen

Para analizar el funcionamiento diferencial de los ítems de Teste Cloze por Opção- EM/U que evalúa la comprensión lectora, participaron 489 estudiantes. De ellos, 206 estudiantes de una institución privada, siendo 79,1% mujeres, edad media de 25,89 años (SD = 6,79). Otros 283 Estudiantes de Enseñanza Media (EM) con una edad media de 15,98 años (SD = 1,27), 52,7% de participantes de las escuelas públicas y el 54,8% de ellas eran mujeres. La prueba se aplicó colectivamente. Se encontró que cuatro elementos tuvieron las mayores tasas de dificultad. Sólo había un producto fuera del rango de adecuado de outfit. Los ítems están de acuerdo en su y es relativamente fácil para estos estudiantes responder a prueba. Doce elementos funcionan diferencialmente para EM y Superior. Se requieren estudios adicionales para aumentar el conocimiento sobre la prueba y idoneidad para la población destinataria.

Keywords: Funcionamiento diferencial del ítem; Comprensión de lectura; Teoría de respuesta al ítem.

Introdução

Ao longo do processo de escolarização, considerado desde o Ensino Fundamental (EF) até a Universidade, espera-se que o estudante esteja capacitado e desenvolva habilidades para compreender a linguagem e aplicar o conhecimento adquirido, dentre outras relacionadas às atividades básicas da linguagem (Gomes & Boruchovitch, 2009; Joly & Paula, 2005). É esperado que, ao final do 4º ano do EF, os alunos tenham desenvolvido o senso crítico para a leitura. Assim, ao longo dos próximos anos escolares, a competência dos estudantes deve ser aperfeiçoada. Para que o estudante do Ensino Médio (EM) não tenha dificuldade em prosseguir para o nível universitário, os Parâmetros Curriculares Nacionais (PCN – MEC, 2000) propõem a capacitação no que se refere à análise, compreensão, interpretação e aplicação da linguagem, tanto no aspecto receptivo quanto expressivo. Por isso, no nível universitário espera-se que o aluno consiga aprender de forma independente, aplique seu conhecimento na solução dos problemas com os quais se depara, sintetize a informação do texto com a sua experiência, assuma uma atividade avaliativa, revelando uma postura mais crítica diante do processo de aquisição e comunicação da informação de modo eficaz (Cabral, 2006; Joly & Paula, 2005).

A aprendizagem e construção do conhecimento ocorrem por meio da leitura e de sua compreensão. Compreender um texto significa apreender o sentido dado pelo autor e envolve, muitas vezes, o ato de transformar, relacionar e aplicar o conhecimento prévio e aquele adquirido no momento da leitura, assim como a utilização de pensamento crítico, elaboração de predições, entre outras capacidades (Duke & Pearson, 2002; Gomes & Boruchovitch, 2009; Joly & Lomônaco, 2003; Oliveira, Cantalice & Freitas, 2009). Logo, a compreensão da leitura assume papel imprescindível na construção do conhecimento. O entendimento das informações lidas é resultado de um esforço cognitivo direcionado ao modo como a tarefa de compreender é realizada, tanto do ponto de vista do leitor quanto da própria tarefa (Brandão & Spinillo, 1998; Gonçalves, 2008; Joly & Lomônaco, 2003).

No que se refere à forma de mensuração do nível de compreensão em leitura, a técnica de Cloze vem sendo comumente utilizada e, atualmente, determinada como adequada tanto para o diagnóstico como para a intervenção nessa área (Joly, 2006, 2009b, Joly & Marini, 2006). Essa técnica surgiu em 1953, proposta por W. Taylor e consiste, originalmente, na seleção de um texto de aproximadamente 300 palavras, no qual se omite uma palavra a cada cinco, substituindo-a por um espaço proporcional ao seu tamanho. A necessidade de pensar e escolher uma palavra que seja adequada para completar a lacuna, a fim de dar sentido ao texto, exige que o respondente o compreenda. Posteriormente, algumas adaptações do formato original dessa técnica foram sendo estudadas, como fez Riley (1986) quando adaptou para uma versão progressiva que consistia na inserção gradativa de dificuldades na tarefa de compreensão para que o aluno construísse o significado do texto.

Joly (2006; 2009a; 2009b) também fez uma adaptação visando determinar níveis de dificuldade diferenciados utilizando a técnica de Cloze por meio da organização do texto a partir de critérios

específicos relativos ao número de palavras, omissões de vocábulos, tamanho de lacunas e opções de resposta. Essa variação foi denominada de Sistema Orientado de Cloze (SOC) e pode ser utilizada de modo estratégico a fim de ampliar as habilidades requeridas na compreensão de um texto. O SOC é organizado a partir de um ou mais critérios específicos, que, por sua vez, determinarão a dificuldade do texto. Nesse tipo de teste, cada palavra omitida se configura como um item.

Considera-se que o indivíduo que responde um teste de Cloze se caracteriza como um processador de textos que faz inferências e analogias resultantes da utilização de seu conhecimento prévio (Abraham & Chapelle, 1992). Assim, cada item terá sua dificuldade, pois cada respondente tem suas características específicas nas habilidades necessárias para essa atividade. É importante destacar que cada respondente adotará determinadas estratégias para realizar essa atividade, de acordo com suas características e necessidades.

No que se refere às avaliações já realizadas, resultados de pesquisas, inclusive em nível nacional, que estudaram a compreensão em leitura de alunos do Ensino Fundamental (Carvalho, 2006; INEP, 2006a; Joly & Nicolau, 2005) e Médio (INEP, 2006b), têm mostrado que o nível desejável de compreensão em leitura para os diversos graus de escolarização não está sendo alcançado. À guisa de exemplo, um estudo desenvolvido por Joly, Santos e Marini (2006) com alunos de EM de escolas públicas e privadas, evidenciou a dificuldade dos alunos quanto ao entendimento dos textos (adequados ao nível de escolaridade). Foi verificado que os estudantes utilizam mais estratégias de solução de problemas na leitura, o que denota ações empreendidas quando há dificuldade para entendimento dos textos.

Alunos de um curso pré-vestibular foram participantes do estudo de Oliveira, Cantalice e Freitas (2009) e responderam um teste de Cloze cujo desempenho máximo possível era de 46 pontos. A pontuação máxima verificada foi de 25 pontos, sendo que a média de acertos desses estudantes foi igual a 16,7 pontos, ou seja, menor do que a metade do escore possível. Assim, revelaram possuir sérias dificuldades de leitura e compreensão de textos. Tendo em vista que todos esses alunos almejam matricular-se em uma universidade e, considerando a etapa de conclusão do EM, o panorama relativo à habilidade de compreender textos mostra-se altamente desfavorável.

Acrescente-se a esses resultados os dados de pesquisas com universitários, tanto ingressantes quanto concluintes. Entre eles, os encontrados por Joly e Paula (2005) em um estudo objetivando conhecer o nível de compreensão em leitura de universitários ingressantes, de diversas áreas de conhecimento, por meio de um teste de Cloze. Os resultados revelaram que a maioria dos estudantes teve compreensão empobrecida do texto, ou seja, os estudantes não eram hábeis para compreender um texto da forma exigida nesse nível de escolaridade. Outros estudos que também avaliaram a compreensão em leitura de universitários encontraram resultados nessa direção (e.g. Joly & Guerra, 2004; Joly & Paula, 2005; Oliveira, 2003; Santos et al., 2006). A partir desses estudos e considerando outros (por exemplo, Carvalho, 2006; Joly & Nicolau, 2005) que avaliaram a compreensão em leitura de estudantes de níveis de ensino anteriores e chegaram a resultados aquém do esperado, conclui-se que os alunos chegam à universidade sem as habilidades necessárias para atender às exigências desse nível escolar no que se refere à compreensão em leitura.

A partir desse panorama, o presente estudo tem por objetivo analisar o nível de dificuldade dos itens de um instrumento de avaliação da compreensão em leitura, o Teste Cloze por Opção - EM/U, aplicado a estudantes do EM e de nível universitário, considerando também a análise de funcionamento diferencial (DIF) de seus itens em função do nível acadêmico/escolar. Pesquisas já realizadas com o Teste Cloze por Opção - EM/U serão posteriormente descritas possibilitando uma análise crítica do desempenho dos alunos desses níveis de escolaridade especificamente quando a compreensão em leitura foi aferida por meio desse teste. Alunos do EM de escolas públicas e particulares foram participantes do estudo de Marini (2006) e responderam ao Teste Cloze por Opção, baseado no SOC. A média de acertos foi de 33,1 pontos, com escores variando de 1 a 40 (máximo possível). Dessa amostra, 5% atingiu o escore máximo possível.

Dias (2008) também utilizou o Teste Cloze por Opção como um dos instrumentos de seu estudo com universitários. Verificou-se que a pontuação média foi igual a 32,7 pontos. Além disso, 49,9% do total da amostra obteve pontuação no Cloze entre 35 a 38 pontos, evidenciando que quase a metade dos participantes atingiu pontuações próximas ao escore máximo dessa prova. Um destaque, ao lado disso, é que nenhum participante obteve o escore máximo possível. Embora tenha sido verificado um nível satisfatório de compreensão em leitura por parte de aproximadamente metade da amostra universitária do estudo de Dias (2008), há que se considerar a outra metade que não alcançou tal nível. Sublinhamos que o estudante universitário recebe, via de regra, grande demanda de leitura para adquirir o conhecimento específico da área estudada. Martins (2008) pesquisou a compreensão em leitura de universitários por meio do mesmo Teste Cloze por Opção e verificou que os participantes compreenderam 81% do texto. O desempenho médio foi igual a 32,4 pontos, sendo que a pontuação

máxima foi atingida por alguns dos participantes. Com a aplicação da TRI, notou-se que o texto tinha um nível de dificuldade baixo para os alunos que compuseram sua amostra.

Por meio da análise dos resultados desses estudos, nota-se que as pontuações médias das amostras de EM e universitário foram bastante similares. Ao mesmo tempo, considerando os dados de diversos estudos realizados com universitários, que indicam desempenho aquém do esperado (Joly & Guerra, 2004; Joly & Paula, 2005; Oliveira, 2003; Santos, et al., 2006), pode-se supor que não há avanço muito grande relativo à compreensão em leitura na transição do EM para o universitário. Essa constatação, entretanto, deve ser mais especificamente investigada a partir de: análises do nível de dificuldade do item, via Teoria de Resposta ao Item; análise de um item para verificar se o fato dele funcionar de forma diferente para pessoas que, agrupadas em função de alguma variável e tendo o mesmo nível de habilidade na área avaliada e, conseqüentemente, a mesma chance de acerto, fornecem resposta contrária à esperada, evidencia que ele precisa ser reajustado. Isso porque o desempenho em um teste baseado em respostas de itens que funcionam de maneira diferente para um grupo de pessoas pode influenciar a tomada de decisões a partir desse resultado. No caso de instrumentos que buscam aferir o nível de compreensão em leitura é importante que o diagnóstico e as inferências advindas da avaliação sejam realizados com mais segurança e confiabilidade.

Privilegiar um grupo em detrimento de outro que, mesmo tendo características diferentes, possui habilidade para desempenhar-se de maneira idêntica ao primeiro, não deve ser um objetivo de nenhuma avaliação ou testagem psicológica, pois seria um teste (ou conseqüente diagnóstico) injusto e inadequado (Andriola, 2001). Assim, idealmente, pessoas com nível de habilidade maior devem ter, também, maior desempenho no teste em questão do que aquelas com menor habilidade e vice-versa (Sisto, 2006). Cabe ressaltar que diferenças nas respostas de um item por parte de diferentes grupos podem indicar reais diferenças na variável mensurada. Nesse caso, teria sido verificado o impacto do item (Pasquali, 2007). A análise do funcionamento diferencial dos itens (DIF) feita por meio do programa *Winsteps* é apenas um dos vários métodos para se verificar esse tipo de dado (Andriola, 2001). Estudos dessa natureza contribuem para a validade do teste por garantirem que a medida obtida não se encontra afetada por características étnico-culturais dos grupos submetidos à avaliação, entre outras, como a própria redação do item (Andriola, 2001; Sisto, 2006).

Por meio do *Winsteps*, a análise de DIF implica verificar os parâmetros relacionados ao ajustamento dos itens, como o índice de *infit* e *outfit*. Um item estará ajustado quando as respostas observadas não são estatisticamente diferentes das esperadas ou modeladas. Esses índices indicam o quanto as pessoas responderam àquele item especificamente em consonância com as respostas dadas aos outros itens que compõem o teste estudado (Nunes, 2009; Wright & Stone, 1999). Discrepâncias ou desajustes ocorrem, portanto, quando o respondente acerta um item que é muito difícil para ele ou erra um item que seria fácil pra ele. Assim, desajustes que ocorram perto do nível de habilidade do sujeito indicam índice de *infit* e revelam padrões de resposta inesperados, sugerindo que pessoas hábeis para acertar o item o erraram ou o oposto. Discrepâncias que ocorram longe do nível de habilidade do sujeito configuram-se como *outfit* e revelam problemas estruturais nos itens (Nunes, 2009). Alguns autores têm delimitado o intervalo adequado de valores para os parâmetros de ajustamento. Segundo Linacre (2006), valores de *infit* e *outfit* entre 0,70 e 1,50 são aceitáveis. Para definir se um item possui DIF, Draba (1977) propôs valores de *t* maiores de 2,4.

O estudo do DIF, portanto, torna-se importante para verificação de medidas do item considerando variáveis de caráter social e psicométrico. No entanto, é necessário ressaltar que a exclusão de itens com base somente na análise estatística, por exemplo, a de funcionamento diferencial (DIF), deve ser uma decisão norteada também por avaliações teóricas, feitas por especialistas da área estudada. Além disso, é necessária a replicação dos estudos para que se confirmem os resultados, já que há a possibilidade, por exemplo, de o DIF ser acidental (Pasquali, 2007).

O objetivo do presente estudo é avaliar o Teste Cloze por Opção EM/U, uma prova de compreensão em leitura baseada na técnica de Cloze e aplicada a alunos do Ensino Médio e universitário, por meio da Teoria de Resposta ao Item, e verificar se há itens com funcionamento diferencial em função do nível de ensino em busca de evidências de validade para o Teste Cloze por Opção EM/U. Atrelado a esse objetivo, está a hipótese de que não serão encontrados itens com funcionamento diferencial já que, por meio dos estudos já realizados (e.g. Dias, 2008; Joly, Santos & Marini, 2006; Marini, 2006; Martins, 2008; Oliveira, Cantalice & Freitas, 2009), percebem-se níveis de compreensão em leitura de alunos do EM e universitários bastante semelhantes.

Método

Participantes

A coleta de dados foi realizada em instituições de ensino localizadas no interior do estado de São Paulo. Participaram 489 estudantes, sendo 206 universitários e 283 alunos do Ensino Médio. Dentre os universitários, 23,3% ($N = 48$) deles estavam no 1º ano da graduação, 21,4% ($N = 44$) no segundo ano, 16,5% ($N = 34$) no terceiro, 18% ($N = 37$) no quarto e 20,9% ($N = 43$) no quinto ano. Já quanto aos do EM, 33,2% ($N = 94$) estavam no 1º ano, 33,6% ($N = 95$) no 2º ano e 33,2% ($N = 94$) no terceiro.

Tabela 1 – Caracterização dos participantes em relação ao gênero, tipo de ensino e período de estudo.

Variáveis de caracterização		Amostra EM		Amostra universitários	
		%	N	%	N
Gênero	Feminino	54,8	155	79,1	163
	Masculino	45,2	128	20,9	43
Ensino	Público	52,7	149	-	-
	Privado	47,3	134	100	206
Período de estudo	Diurno	72,1	204	17,5	36
	Noturno	27,9	79	82,5	170

Instrumento

Teste Cloze por Opção – EM /U (Joly, 2005)

É composto por um texto do tipo editorial ao qual foi aplicada a Técnica de Cloze (Taylor, 1953), com alternativas de múltipla escolha. Foi omitido todo quinto vocábulo num texto de 300 palavras. As palavras omitidas foram substituídas por lacunas, sendo todas do mesmo tamanho. No total, há 40 lacunas. Os sujeitos escolhem, entre as cinco alternativas de múltipla escolha, a palavra que preenche corretamente cada lacuna, dando sentido à frase. As alternativas de múltipla escolha possuem uma única resposta correta, uma palavra da mesma classe gramatical da palavra que foi omitida, duas palavras relacionadas à classe gramatical que foi omitida e uma palavra de classe gramatical diferente da omitida.

São consideradas corretas as palavras que completam a lacuna de forma exatamente igual ao texto original. As respostas em branco são computadas como erro, para cada acerto é atribuído um ponto. O rendimento dos sujeitos é registrado considerando acertos e erros. A validade do teste foi obtida em estudos exploratórios (Dias, 2008; Joly & Martins, 2006). A precisão do instrumento é de 0,90, obtida utilizando-se o teste Alfa de Cronbach e de 0,81 pelo coeficiente de Kuder-Richardson (KR-21). No estudo de Dias (2008), os resultados indicaram que houve evidências de validade convergente com outro instrumento que mede compreensão, com correlação significativa ($r=0,53$; $p<0,001$). A análise fatorial com informação completa revelou predominância unidimensional para os itens, com explicação de 24,80% da variância.

Procedimento

Este estudo seguiu os padrões éticos de pesquisa. A aplicação ocorreu de forma coletiva, por sala de aula, tanto nas escolas do Ensino Médio quanto na Universidade. O aplicador dava as instruções e a realização da atividade teve duração média de 30 minutos por sala, não tendo sido estabelecido tempo limite para responder.

Resultados

Inicialmente, foi realizada a análise dos itens do Teste Cloze por Opção-EM/U de Joly (2005) por meio do modelo logístico de um parâmetro (modelo de Rasch) baseado na Teoria de Resposta ao Item (TRI) para, posteriormente, fazer-se a análise de DIF. A Tabela 2 apresenta a análise do instrumento considerando a estimação do parâmetro de dificuldade dos itens pelo Rasch.

Tabela 2 - Índices estatísticos do ajuste dos itens, dificuldades dos itens, infits, outfit.

Item	Modelo de Rasch			Item	Modelo de Rasch		
	<i>b</i>	<i>infit</i>	<i>outfit</i>		<i>b</i>	<i>infit</i>	<i>outfit</i>
1	-2,07	1,16	1,34	21	-0,60	1,06	1,02
2	-1,80	1,00	1,26	22	-1,38	0,92	0,60
3	-1,27	0,89	0,73	23	1,27	1,09	1,14
4	-0,78	1,00	0,72	24	-1,57	0,86	0,65
5	-0,11	0,91	0,85	25	-0,56	0,92	0,59
6	1,28	1,04	0,92	26	0,79	0,96	1,04
7	-1,51	0,78	0,43	27	2,07	1,07	1,09
8	0,41	1,07	1,08	28	-0,26	1,03	1,38
9	0,46	1,13	1,43	29	-0,12	1,01	1,05
10	-1,50	0,84	0,64	30	-0,36	0,98	0,89
11	-2,28	0,83	0,23	31	0,73	1,02	0,95
12	2,74	1,06	1,07	32	-0,91	0,78	0,58
13	0,57	1,01	0,92	33	3,39	1,25	1,58
14	0,28	0,91	0,67	34	1,51	1,02	1,01
15	-0,13	1,00	1,03	35	3,32	1,11	1,38
16	1,30	0,92	0,74	36	-0,83	1,13	0,87
17	0,08	1,05	1,25	37	0,55	0,87	0,69
18	0,60	0,95	0,83	38	-0,31	0,86	0,46
19	-0,59	0,94	0,98	39	-0,95	0,94	0,71
20	-1,50	0,75	0,47	40	0,03	1,10	1,07

Observa-se na Tabela 2 que os itens com maiores índices de dificuldade foram os de número 33 (advérbio), 35 (substantivo), 12 (conjunção aditiva) e 28 (conjunção causal). Pode-se supor, por um lado, que as classes gramaticais dessas palavras possam ter aumentado a dificuldade dos itens por exigirem muita habilidade para uso adequado com a língua, habilidade que não se observa no desempenho dos estudantes de EM (INEP, 2006b) e Superior (Joly & Guerra, 2004; Joly & Paula, 2005; Oliveira, 2003; Santos, et al., 2006), e, por outro lado, que talvez as opções colocadas para esses itens dificultem a escolha dos alunos.

Para nenhum dos itens foi verificado índice de *infit* fora do limite crítico, sendo o intervalo de 0,70 a 1,50 considerado como critério de bom ajuste (Linacre, 2006). Assim, constata-se que o conjunto de itens do Teste Cloze por Opção EM/U (Joly, 2005) não revelou discrepâncias na região distante das habilidades estimadas dos sujeitos e que, então, está ajustado do ponto de vista estrutural. A média do índice *infit* foi igual a 0,98 ($DP = 0,11$), sendo que os valores variaram de 0,75 a 1,25. Já para o *outfit* a média foi de 0,91 ($DP = 0,30$) com variação de 0,23 a 1,58. Esses índices indicam, no geral, que a maioria das pessoas apresentou o padrão esperado de resposta de acordo com Linacre (2006).

Quando se considera o índice de *outfit* verifica-se que somente o item 33 esteve fora do limite adequado. Esse dado sugere que as respostas dadas a esse item não estavam dentro dos padrões esperados, considerando a habilidade estimada dos respondentes. Em análise específica do item 33, verificou-se que 57% dos estudantes da amostra assinalaram uma opção que pertencia à mesma classe gramatical da resposta correta, mas que era errada. Nota-se, por meio dessa análise, que esse item apresenta uma opção de resposta que é o antônimo da resposta correta (no caso, mais - menos) e que a maior parte dos alunos escolheu o antônimo, revelando que consideraram o nível mais micro do que o macrotextual, já que atribuiu sentido à frase especificamente, o qual não condizia com o contexto (Abraham & Chapelle, 1992). Vale citar que 39% da amostra respondeu corretamente esse item. De todo modo, pode-se inferir que boa parte da amostra pode não ter aplicado o conhecimento sobre as atividades básicas de compreensão de uma leitura, conforme indicam Cabral (2006), Gomes e Boruchovitch (2009) e Joly e Paula (2005), como a análise do contexto, ao menos com relação especificamente a esse item, dada sua dificuldade, por se tratar de advérbio.

Outro critério de análise para os itens foi verificar como se distribuem em função das habilidades dos estudantes participantes, independentemente do nível de ensino que cursavam. O mapa de itens do desempenho no Teste de Cloze por Opção EM/U está apresentado na Figura 1, abaixo.

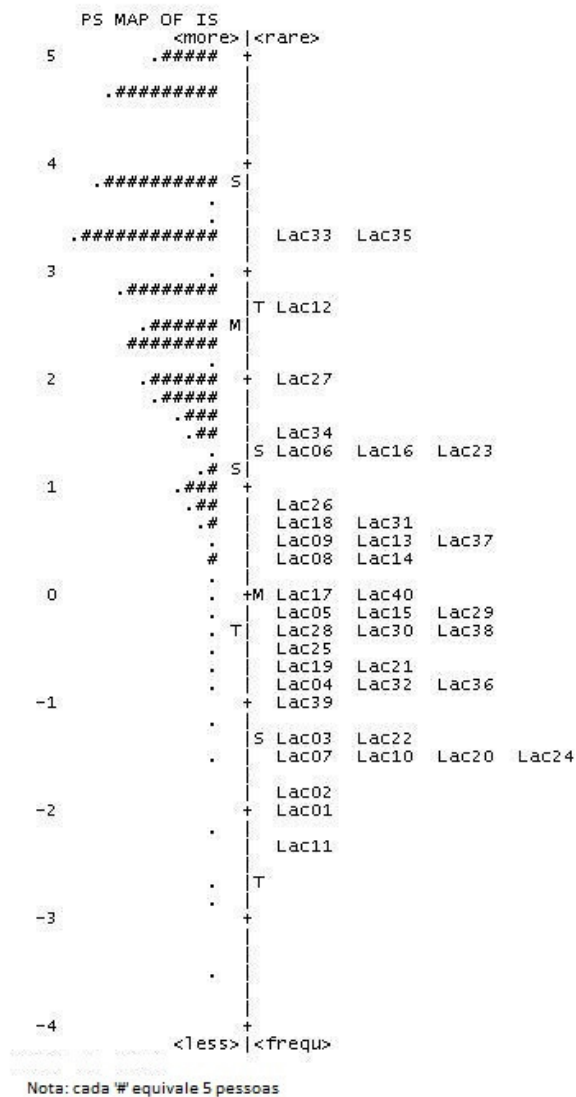


Figura 1 – Mapa de pessoas e itens pela TRI, considerando-se o theta das pessoas e o parâmetro b dos itens.

A Figura 1 oferece as informações sobre o nível da habilidade dos sujeitos (*theta*) e o nível (*b*) de dificuldade dos itens numa mesma escala de medida (*logit*). É possível notar que a habilidade das pessoas foi estimada no intervalo de -3,52 a +5,97, com média 2,71 (*DP* = 1,57). A dificuldade dos itens variou de -2,28 a +3,39, com média 0,00 (*DP* = 1,36). Nota-se, também, que a distribuição dos participantes, considerando a habilidade para responder a prova, não está na mesma amplitude do conjunto de itens, de acordo com as relativas dificuldades estimadas.

Foi observado que 25 estudantes (5,1%) atingiram o escore máximo no instrumento, sendo que 11 deles frequentavam o Ensino Médio e 14 o Ensino Superior. Constata-se que o nível independente de desempenho em medidas de compreensão de leitura, no qual os estudantes revelam condições básicas para ler, é atingido pela minoria da amostra. O resultado é corroborado por Joly e Paula (2005), Joly, Santos e Marini (2006), Oliveira, Cantalice e Freitas (2009). A despeito disso, considerando a habilidade média da amostra e a dificuldade média do teste, constata-se que é relativamente fácil para os estudantes que compuseram a amostra deste estudo responder ao Teste Cloze por Opção EM/U, já que o nível médio de *theta* dos respondentes tendeu a ser maior que o nível médio de *b* dos itens. O Sistema Orientado de Cloze, nessa perspectiva, pode ser aplicado de modo a ajustar a dificuldade do teste (Joly, 2006, 2009a, 2009b). Assim, pode-se inferir que, embora o texto seja adequado para o nível de ensino, as opções de resposta, em sua maioria, contribuíram para a relativa facilidade da prova para os alunos que compuseram esta amostra, exceção feita para os itens 33, 35, 12 e 28. Contudo, ressalta-se a

homogeneidade da amostra de cada nível de ensino, sendo indicada uma amostra mais ampla e representativa para que os resultados sejam conclusivos (Maroco, 2007).

Após a avaliação dessas propriedades do Teste Cloze por Opção-EM/U, realizou-se o estudo do DIF considerando os tipos de ensino (Médio e Superior). O DIF foi calculado por meio do programa *Winsteps* que, usando o teste *t*, verifica se o resíduo resultante da comparação entre os desempenhos dos estudantes de Ensino Médio e dos universitários pode ser atribuído ao acaso ou não. Utilizou-se o valor crítico igual ou superior a 2,4 como referência para atribuição de significância estatística, pois, de acordo com Draba (1977), esse é o índice adequado quando se está analisando o funcionamento diferencial de mais de 20 itens.

Tabela 3 - DIF por nível de escolaridade, mudança e valores de t por item.

Item	Nível de escolaridade	DIF	Nível de escolaridade	DIF	Mudança	t
1	EM	-3,72	ES	-0,81	-2,91	-3,55
2	EM	-2,00	ES	-1,45	-0,54	-0,93
3	EM	-1,19	ES	-1,46	0,27	0,50
4	EM	-0,77	ES	-0,80	0,03	0,06
5	EM	-0,14	ES	-0,07	-0,06	-0,18
6	EM	1,23	ES	1,36	-0,12	-0,49
7	EM	-1,64	ES	-1,27	-0,37	0,53
8	EM	-0,18	ES	1,17	-1,35	-4,67
9	EM	0,23	ES	0,81	-0,58	-2,01
10	EM	-1,63	ES	-1,27	-0,37	-0,69
11	EM	-2,49	ES	-1,94	-0,55	-0,79
12	EM	2,70	ES	2,80	-0,09	-0,43
13	EM	0,26	ES	1,03	0,77	-2,74
14	EM	0,13	ES	0,53	-0,40	-1,31
15	EM	-0,06	ES	-0,28	0,22	0,61
16	EM	1,57	ES	0,80	0,77	2,93
17	EM	0,21	ES	-0,19	0,40	1,15
18	EM	0,53	ES	0,72	-0,18	-0,64
19	EM	-0,61	ES	-0,55	-0,06	-0,15
20	EM	-1,53	ES	-1,46	-0,07	-0,13
21	EM	-1,11	ES	0,11	-1,23	-3,23
22	EM	-1,53	ES	-1,11	-0,42	-0,82
23	EM	1,75	ES	0,18	1,57	5,31
24	EM	-1,44	ES	-1,91	0,47	0,76
25	EM	-0,50	ES	-0,68	0,18	0,43
26	EM	0,36	ES	1,38	-1,02	-3,85
27	EM	2,04	ES	2,11	-0,08	-0,35
28	EM	-0,17	ES	-0,47	0,30	0,79
29	EM	-0,50	ES	0,43	-0,93	-2,81
30	EM	-0,18	ES	-0,81	0,64	1,53
31	EM	0,72	ES	0,75	-0,03	-0,11
32	EM	-0,90	ES	-0,95	0,05	0,10
33	EM	3,04	ES	3,88	-0,85	-3,72
34	EM	1,51	ES	1,52	-0,02	-0,06
35	EM	4,12	ES	2,42	1,70	7,09
36	EM	-0,50	ES	-1,93	1,43	2,46
37	EM	0,89	ES	-0,33	1,23	3,50
38	EM	-0,16	ES	-0,68	0,52	1,28

Item	Nível de escolaridade	DIF	Nível de escolaridade	DIF	Mudança	t
39	EM	-1,10	ES	-0,68	-0,42	-0,96
40	EM	0,24	ES	-0,45	0,69	1,87

A Tabela 3 indica que 12, dentre os 40 itens que compõem o instrumento, mostraram-se significativamente diferentes em função do nível de ensino no qual os estudantes desta amostra estão inseridos. Esses itens foram os de número 01, 08, 13, 16, 21, 23, 26, 29, 33, 35, 36 e 37. É possível notar que, entre os 12 itens que apresentaram DIF, 7 indicam dificuldade maior em responder para estudantes do Ensino Superior, enquanto 5 sugerem dificuldade maior para os do Ensino Médio. Desse modo, verifica-se certo equilíbrio nos vieses observados: entre os dois níveis de ensino, os universitários tiveram menor favorecimento do que os alunos do Ensino Médio. Isso revela a possibilidade de um processamento cognitivo diferenciado para EM e Superior quanto à compreensão de textos cuja temática seja genérica, implicando em estratégias diferentes para recuperar conhecimento prévio e utilizá-lo para completar os itens da prova (Abraham & Chapelle, 1992; Duke & Pearson, 2002).

Considerações Finais

Fazem-se necessários novos estudos para verificar se esses itens funcionam diferentemente em função do nível de ensino, dada a limitação do presente estudo no que se refere à distribuição da amostra em função da habilidade dos sujeitos. Ainda assim, considerando que os índices de ajuste dos itens estiveram dentro do esperado e que isso, conseqüentemente, indica presença mínima de respostas inesperadas, evidencia-se validade para o Teste Cloze por Opção EM/U pela análise da TRI a partir do Modelo de Rasch (Andriola, 2001; Linacre, 2006; Draba, 1977; Nunes, 2009; Wrigth & Stone, 1999).

Dada a relativa facilidade em responder ao teste, sugerem-se também estudos com alunos do Ensino Fundamental a fim de verificar se o nível de dificuldade da prova pode ser adequado para essa etapa de escolaridade. Igualmente, estudos com o mesmo instrumento, mas adequando seu nível de dificuldade a partir dos critérios propostos pelo SOC, como, por exemplo, sem opções de resposta. Estudos que considerem o mesmo texto organizado de diferentes modos, a partir desses critérios, e que verifiquem o nível de dificuldade dessas versões, comparando-as, podem acrescentar parâmetros de análise mais seguros para diagnosticar o nível da compreensão de leitura dos alunos universitários e de Ensino Médio (Pasquali, 2007). Também, é conveniente que estudos futuros objetivem realizar a análise de DIF considerando outras variáveis, como gênero, período do dia em que estuda, nível sócio-econômico, tipo de escola frequentada, faixa etária, entre outras. Além disso, é necessária a exploração das causas dessas diferenças para que a decisão de exclusão de algum item ou de adaptação semântica do texto seja realizada com mais segurança.

Referências

- Abraham, R. G. & Chapelle, C. A. (1992). The meaning of Cloze test scores: an item difficulty perspective. *The Modern Language Journal*, 76, 468-479.
- Andriola, W. B. (2001). Descrição dos principais métodos para detectar o Funcionamento Diferencial dos Itens (DIF). *Psicologia, Reflexão e Crítica*, 14(3), 643-652.
- Brandão, A. C. P. & Spinillo, A. G. (1998). Aspectos gerais e específicos na compreensão de textos. *Psicologia, Reflexão e Crítica*, 11(2), 253-272.
- Cabral, A. P. S. (2006). Fórmulas de intervenção nos domínios da compreensão em leitura e escrita no ensino superior. In: M. C. R. A. Joly & C. Vectore, *Questões de pesquisa e práticas em Psicologia Escolar*. (pp. 61-83). São Paulo: Casa do Psicólogo.

Carvalho, M. R. (2006). *Estratégias metacognitivas de leitura utilizadas de 2ª a 4ª série do Ensino Fundamental*. Dissertação de Mestrado, Universidade São Francisco, Itatiba, São Paulo.

Dias, A. S. (2008). *Evidências de validade de uma prova de compreensão em leitura em estatística*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Draba, R. E. (1977). *The identification and interpretation of item Bias* (MESA Memorandum No. 25). Chicago: The University of Chicago. Acessado em setembro de 2010. Disponível em <http://www.rasch.org/memo25.htm>.

Duke, N. K. & Pearson, D. (2002). Effective Practices for Developing Reading Comprehension. In: A. E. Farstrup & S. J. Samuels (Eds.). *What Research Has To Say About Reading Instruction* (3ª ed., pp. 205-242). Newark, DE: International Reading Association.

Gonçalves, S. (2008). Aprender a ler e compreensão do texto: processos cognitivos e estratégias de ensino. *Revista Iberoamericana de Educación*, 46, 135-151.

Gomes, M. A. & Boruchovitch, E. (2009). Proficiência em leitura: um panorama da situação. In: A. A. Santos, E. Boruchovitch, & K. L. Oliveira. *Compreensão de leitura: o Cloze como instrumento de diagnóstico e intervenção*. (pp. 23-46) São Paulo: Casa do Psicólogo.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP (2006a). *Qualidade da educação: uma nova leitura do desempenho dos estudantes de 8ª série do Ensino Fundamental*. Dezembro, 2003. [On-line] Recuperado: 16 de out. 2006. Disponível: http://www.inep.gov.br/download/cibec/2003/saeb/qualidade_educ.pdf

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP (2006b). *Qualidade da educação: uma nova leitura do desempenho dos estudantes de 3º ano do Ensino Médio*. Janeiro, 2004. [on-line] Recuperado: 16 de out. 2006. Disponível: http://www.inep.gov.br/download/saeb/2004/qualidade_educacao.pdf

Joly, M. C. R. A. & Guerra, P. B. C. (2004). Compreensão em leitura e barreiras à criatividade: um estudo com universitários ingressantes. *PSICO*, 35, 151-159.

Joly, M. C. R. A. (2005). *Teste de Cloze por opção*. (Relatório Técnico). Universidade São Francisco, Itatiba.

Joly, M. C. R. A. (2006). Cloze Oriented System (COS) in an electronic comprehension program and reading attitude in Brazil. Em A. Mendéz-Villas, B. G. Pereira, J. M. González & J. A. M. González (Ed.), *Current Developments in Technology-Assisted Education* (pp. 1669-1674). Badajoz: INDUGRAFIC.

Joly, M. C. R. A. (2009a). Avaliação da compreensão em leitura pelo Sistema Orientado de Cloze (SOC). *Revista Mexicana de Orientación Educativa*, 17, 1-8.

Joly, M. C. R. A. (2009b). Estudos com o Sistema Orientado de Cloze para o Ensino Fundamental. Em A. A. Santos, E. Boruchovitch & K. L. Oliveira. *O Cloze como instrumento de diagnóstico e intervenção*. (pp. 119-146). São Paulo: Casa do Psicólogo.

Joly, M. C. R. A. & Lomônaco, J. F. B. (2003). Avaliando a compreensão de leitura no Ensino Fundamental: uma comparação entre o instrumento eletrônico e o impresso. *Boletim de Psicologia*, 53(119), 131-147.

Joly, M. C. R. A. & Marini, J. A. S. (2006). Metacognição e Cloze na avaliação de dificuldades em leitura. Em: M. C. R. A. Joly, & C. Vectore, *Questões de pesquisa e práticas em Psicologia Escolar*. (pp. 13-36). São Paulo: Casa do Psicólogo.

Joly, M. C. R. A. & Martins, R. X. (2006). Estudo de validade de uma escala de desempenho em tecnologias para estudantes. *Psicologia Escolar e Educacional*, 10, 41-52.

Joly, M. C. R. A. & Nicolau, A. F. (2005). Avaliação de compreensão em leitura usando Cloze na 4ª série. *Temas sobre desenvolvimento*, 14 (83-84), 14-19.

Joly, M. C. R. A. & Paula, L. M. (2005). Avaliação do uso de estratégias de aprendizagem e a compreensão em leitura de universitários. Em M. C. R. A. Joly, A. A. A. Santos, & F. F. Sisto (orgs.), *Questões do cotidiano universitário*. (pp. 33-58). São Paulo: Casa do Psicólogo.

Joly, M. C. R. A., Santos, L. M. & Marini, J. A. S. (2006). Uso de estratégias de leitura por alunos do Ensino Médio. *Paidéia*, 16 (34), 205-212.

Linacre, J. M. (2006). *A user's guide to Winsteps Ministep: Rasch-Model Computer Programs*. Chicago. ISBN 0-941938-03-4.

Marini, J. A. S. (2006). *Escala de Estratégias Metacognitivas de Leitura para o Ensino Médio: evidências de validade*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, Itatiba, SP.

Maroco, J. (2007). *Análise estatística com o uso do SPSS*. Lisboa: Sílabo Editora.

Martins, R. X. (2008). *Modalidades de ensino e sua relação com habilidades cognitivas e tecnológicas*. Tese de doutorado, Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, Itatiba, SP.

Ministério da Educação (MEC). (2000). *Parâmetros Curriculares Nacionais (Ensino Médio) – Linguagem, códigos e suas tecnologias*. Brasília, DF: Gráfica do Governo.

Nunes, M. F. O. (2009). *Estudos psicométricos da Escala de Auto-eficácia para atividades ocupacionais*. Tese de doutorado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Oliveira, K. L. (2003). *Compreensão de leitura, desempenho acadêmico e avaliação da aprendizagem em universitários*. Dissertação de Mestrado, Programa de Pós-Graduação Stricto Sensu em Psicologia, Universidade São Francisco, Itatiba, São Paulo.

Oliveira, K. L., Cantalice, L. M. & Freitas, F. A. (2009). Leitura e desempenho escolar em alunos do ensino fundamental. In: A. A. A. Santos, E. Boruchovitch, & K. L. Oliveira. *Compreensão de leitura: o Cloze como instrumento de diagnóstico e intervenção*. (pp. 165-185) São Paulo: Casa do Psicólogo.

Pasquali, L. (2007). A função diferencial dos itens e do teste (DIF e DTF). In: L. Pasquali, *Teoria de Resposta ao Item (TRI): teoria, procedimentos e aplicações*. Brasília-DF: Laboratório de Pesquisa em Avaliação e Medida – LabPAM/UnB.

Riley, J. D. (1986). *Progressive Cloze as a remedial technique*. The Reading Teacher, 39 (6), 576-581.

Santos, A. A. A.; Vendramini, C. M. M.; Suehiro, A. C. B.; & Santos, L. A. D. (2006). Leitura compreensiva e utilização de estratégias de aprendizagem em alunos de psicologia. *Estudos de Psicologia* (Campinas), 23 (1), 83-92.

Sisto, F. F. (2006). O funcionamento diferencial dos itens. *Psico-USF*, 11 (1), 35-43.

Taylor, W. L. (1953). Cloze procedure: a new tool for measuring read-ability. *Journalism Quarterly*, 30, 415-433.

Wright, B. D., & Stone, M. H. (1999). *Measurement essentials*. 2ª ed. Delaware: Wide Range.

Submetido em: 03/05/2011

Revisto em: 09/08/2011

Aceito em: 14/08/2011