

## AJUSTE DE UNA PRUEBA DE CLOZE AL MODELO DE RASCH: CALIBRACIÓN, AJUSTE DE LOS ÍTEMS Y ESTUDIO DEL DIF

### THE ADJUSTMENT OF A CLOZE TEST TO THE RASCH MODEL: CALIBRATION, ADJUSTMENT ITEMS AND DIF ANALYSIS

**Katya Luciane Oliveira**

*Universidade Estadual de Londrina, Brasil*

**Fabián Javier Marín Rueda**

*Universidade São Francisco, Brasil*

**Acácia Aparecida Angeli Santos**

*Universidade São Francisco, Brasil*

**Resumen:** Este estudio tuvo como objetivo investigar el ajuste de una prueba de Cloze al modelo de Rasch y evaluar la dificultad de respuesta al ítem en función del sexo de las personas (DIF). Participaron de la investigación 1022 estudiantes universitarios de los estados de San Pablo e Minas Gerais, Brasil. La prueba de Cloze fue aplicada de forma colectiva. El análisis del instrumento mostró un buen ajuste al modelo de Rasch, así como también que todos los ítems fueron respondidos de acuerdo con el patrón esperado, demostrando de esa forma un buen ajuste. En relación al DIF, apenas cinco ítems indicaron diferenciar el sexo de los participantes. Con base en esos datos, se concluyó que hubo un equilibrio en las respuestas dadas por hombres y mujeres.

**Palabras clave:** Teoría de respuesta al ítem, técnica de Cloze, comprensión lectora.

**Resumo:** Este estudo teve por objetivos investigar o ajuste de um Teste de Cloze ao modelo Rasch e avaliar a dificuldade na resposta ao item em razão do sexo das pessoas (DIF). Participaram da pesquisa 1022 estudantes universitários dos estados de São Paulo e Minas Gerais, Brasil. O teste de Cloze foi aplicado de forma coletiva. A análise do instrumento evidenciou um bom ajuste ao modelo Rasch, bem como os itens foram respondidos conforme o padrão esperado, demonstrando um bom ajuste. Quanto ao DIF, apenas cinco itens indicaram diferenciar o sexo. Com base nos dados, identificou-se que houve equilíbrio nas respostas dadas por homens e mulheres.

**Palavras-chave:** Teoria de resposta ao item, técnica de Cloze, compreensão em leitura.

**Abstract:** The objective of the present study were to investigate the goodness of fit of a Cloze test to the Rasch Model, as well as to evaluate the item difficulty in relation to gender (DIF). The sample was composed of 1022 college students of the São Paulo and Minas Gerais States, Brazil. The cloze test was applied collectively. The analysis of the instrument revealed its goodness of fit to the Rasch Model and that the items were responded according to the expected pattern, showing good adjustment, as well. Regarding DIF, only five items differentiated by gender. Taken together, results indicated a balance in the biases for both man and woman.

**Keywords:** Item theory response, Cloze technique, reading comprehension.

## INTRODUCCIÓN

El comportamiento lector comienza informalmente con lecturas cortas y decodificaciones, pero es en la escuela que se constituye el primer campo estructurado de aprendizaje de la lectura (Flippo, 1998). Cuando empieza el proceso de alfabetización muchos alumnos tienen dificultades

de lectura y de comprensión, siendo común en Brasil que muchos niños concluyan la enseñanza primaria sin tener un dominio real de esa habilidad (Gomes & Boruchovitch, 2009; Navas, Pinto & Delisa, 2009).

Anderson y Pearson (1984), Kintsch (1994), Reeve y Jang (2006) y Venesky (1990) están

de acuerdo con el hecho de que, leer no es una habilidad simple, en tanto envuelve la interacción entre procesamientos cognitivos que incluyen el raciocinio verbal. La lectura no consiste sólo en el acto de descifrar los códigos, sino que consiste en la comprensión crítica y personal de lo que fue leído. Apenas los buenos lectores son capaces de atribuir un nuevo significado y reconstruir el conocimiento, a partir de conocimientos previos sobre el asunto.

La lectura es una habilidad que exige comprensión, fluidez, creatividad y capacidad de deducción (Duke & Pearson, 2002; Ruddell, Ruddell & Singer, 2001; Spira, Bracken & Fischel, 2005). Kintsch (1994) destaca que la lectura exige el trabajo dinámico de esquemas mentales que le permiten al lector, por medio del raciocinio lógico, la elaboración de analogías que lo ayudarán a comprender el sentido completo del texto. Spinillo y Mahon (2007) observan que desde muy temprano los alumnos son capaces de realizar inferencias en la lectura y que lo ideal sería que, con el aumento de la edad y la progresión de los años escolares, el alumno se torne más apto para presentar una buena comprensión lectora.

Pese a eso, algunos autores (Cunha & Santos, 2006; Dias & Vendramini, 2008; Oliveira & Santos, 2005; Silva & Witter, 2009, entre otros) afirman que, los estudiantes brasileiros que llegan a la universidad presentan una serie de desfases en relación a la comprensión lectora. A esos estudiantes les falta comprensión lingüística y contextual, lo cual impide que realicen una articulación crítica entre el texto y los contenidos previamente aprendidos. En muchas universidades brasileiras, la capacidad creativa y de deducción está muy por debajo de lo esperado para ese nivel de escolaridad. Esas consideraciones son preocupantes, especialmente cuando se considera que la lectura con comprensión es importantísima en la educación, una vez que es por medio de ella que los conocimientos técnicos son aprendidos, guiando e instrumentalizando la práctica del futuro profesional (Silva & Witter, 2008).

El proceso de comprensión lectora está llamando cada vez más la atención de los estudiosos, sobre todo en lo que se refiere a la posibilidad de evaluar de forma confiable la habilidad de comprensión textual de estudiantes

de diferentes niveles de la educación formal (Oliveira, Boruchovitch & Santos, 2009). El desafío consiste en conseguir llegar a una medida que permita verificar, de forma confiable, de qué forma ocurre la dinámica de la interacción existente entre el lector y el texto.

Según el modelo de Rasch (1960), una prueba válida debe conseguir discriminar tres aspectos. El primero se refiere al hecho de que una persona con una mayor capacidad tenga también una mayor probabilidad de acertar los ítems, que una persona con una capacidad menor (monotonía); el segundo refiere a que, todas las personas tengan más probabilidades de acertar un ítem fácil, que de acertar un ítem difícil (curvas paralelas); finalmente, en el tercer aspecto, los resultados deben ser producto del desempeño de la persona y de la posición del ítem y no deben sufrir influencia del sexo, la etnia u otras características del sujeto evaluado (ausencia de DIF).

Considerando el significado de esos tres aspectos y la necesidad de tener una prueba válida para analizar la comprensión lectora, se decidió utilizar la Teoría de Respuesta al Ítem (TRI) (Hambleton, R. K., Swaminathan, H. & Rogers, H. J., 1991; van der Linden, W. J., & Hambleton, R. K., 1997). en el análisis de los ítems que componen una medida para verificar la comprensión lectora, conocida como prueba de Cloze. La prueba de Cloze (Taylor, 1953) es empleada como un recurso de diagnóstico y consiste en, un texto de aproximadamente 250 palabras en el cual se omiten todos los quintos vocablos, y en el lugar se coloca un espacio de tamaño proporcional a la palabra que fue omitida. Para responder correctamente a la tarea se requiere que la persona evaluada complete el espacio con la palabra que considera más adecuada, al contexto presente en el texto. Para corregir la prueba, se atribuye un punto para los aciertos, que es cuando la persona completa con la palabra exacta el espacio omitido, y cero puntos para los errores.

Al reflexionar sobre la evaluación de la comprensión lectora, Cunha y Campellini (2010), Cunha y Santos (2006) y Oliveira (2011) consideran que las características del lector, poseen la misma importancia que las características del texto. De esa forma, el resultado de un instrumento con ítems heterogéneos, que utilice la

suma de los aciertos y cuyo producto -de esa suma en el grupo- sea un promedio de aciertos (Teoría Clásica), debe ser visto con cuidado, pues en ese caso específico, podemos entender que no se están respetando las particularidades de cada uno de los ítems que componen un texto, así como tampoco se respetan las características, en función de las habilidades de las personas que respondieron al instrumento.

Otro aspecto, destacado por Santos, Primi, Taxa y Vendramini (2002), Oliveira, Cantalice y Freitas (2009) y Oliveira y Santos (2009) es que los espacios que componen la prueba de Cloze presentan diferentes niveles de dificultad. Esos autores confirman lo descrito por Abraham y Chapelle (1992), que afirmaron que el Cloze presenta ítems más fáciles de ser completados e ítems más difíciles. Además de eso, personas con una habilidad mayor, conseguirían completar tanto los ítems más fáciles como los ítems más difíciles, pues estos últimos son aquellos que exigen la utilización de los procesamientos de abstracción, contextualización y comprensión lingüística (Nicholson, 1999).

De esa forma, al emplearse la TRI en el análisis de la prueba de Cloze, será posible investigar el grado de dificultad de los ítems. Además, según autores como Araújo, Andrade y Bortolotti (2009), Embretson y Reise (2000) y Hambleton, Swaminathan y Hogers (1991), otra ventaja sería entender la razón por la cual algunos ítems tienen una mayor dificultad de acierto, así como también será posible hacer deducciones sobre la habilidad de la persona. Así, el punto de partida es admitir que, ítems heterogéneos no pueden tener el mismo peso y que cada uno de ellos representa la manifestación del rasgo latente (Primi, 2004; van der Linden & Hambleton, 1987).

Entre los modelos propuestos por la TRI (Andrade, Tavares & Valle, 2000) se encuentra el de Rasch (1960). Este análisis sugiere que el acierto de un ítem puede ser explicado sólo en función de la habilidad de la persona y de la dificultad del ítem. Con eso no se daría lo que Holland y Thayer (1988) denominan funcionamiento diferencial del ítem (DIF), que ocurre cuando personas con igual rasgo presentan probabilidades diferentes de acertar el ítem. Lo que se busca en el análisis es un DIF uniforme, en el cual no haya ninguna influencia en la respuesta

por el simple hecho de la persona pertenecer o no a algún grupo específico, como es el caso del sexo (hombres y mujeres).

Hoskins, Newstead y Dennis (1997), consideran que hombres y mujeres pueden presentar diferencias significativas en diferentes habilidades académicas. Específicamente, con foco en la comprensión lectora evaluada por la prueba de Cloze, Boruchovitch, Santos y Oliveira (2007), Oliveira, Santos y Primi (2003), Oliveira y Santos (2005, 2006) y Silva y Santos (2004) verificaron que estudiantes universitarios del sexo femenino, tienen un mejor desempeño que estudiantes del sexo masculino en la prueba de Cloze. Tales investigaciones utilizaron un texto titulado "*Desentendimento*" (Verísimo, 1995), que fue desarrollado especialmente para estudiantes universitarios. De esa forma, se cuestiona el hecho de que la prueba de Cloze podría presentar algún sesgo en su medida, siendo posible que favorezca y privilegie al sexo femenino.

Una justificación para la necesidad de verificar el DIF en la referida prueba de Cloze, es por el hecho de que Santos et al. (2002) también utilizaron la TRI en ese texto (*Desentendimento*), pero en la ocasión, se detuvieron sólo en el análisis del impacto, verificando que los universitarios acertaron más palabras que tenían su sentido relacionado a palabras contiguas, como preposiciones, artículos y sustantivos, siendo que los verbos y los pronombres fueron los ítems más difíciles de ser acertados. Como en ese estudio no se realizaron análisis del DIF, este estudio quiere verificar posibles diferencias entre hombres y mujeres, utilizando la misma prueba de Cloze.

Más recientemente Oliveira y Santos (2009), al realizar un análisis exploratorio de la dificultad de los ítems de la prueba de Cloze '*Desentendimento*', mostraron la misma tendencia de acierto en el Cloze o sea, palabras como preposiciones y artículos fueron las más fáciles de ser acertadas. Pese a que el estudio tuvo un carácter exploratorio y descriptivo, nuevamente se evidenció que los ítems eran heterogéneos y específicos en su dificultad.

Con base en la presentación realizada, este estudio tiene como objetivo investigar el ajuste de una prueba de Cloze (texto *Desentendimento*) al modelo de Rasch, así como

verificar si los ítems demuestran un sesgo asociado al sexo. La necesidad de evaluar la habilidad de comprensión lectora en función del sexo se justifica, pues tiene como objetivo verificar la neutralidad de ese tipo de sesgo. De esa forma, el modelo logístico de Rasch permitirá la realización de una evaluación de la dificultad del ítem y de la habilidad de la persona. Además de eso, será realizado un análisis del DIF en función del sexo de las personas.

## 1. MÉTODO

### 1.1. Participantes

Participaron de la investigación 1022 estudiantes universitarios de los cursos de psicología (47,7%,  $n=488$ ), enfermería (12,2%,  $n=125$ ), matemáticas (12,0%,  $n=123$ ), odontología (10%,  $n=102$ ), administración de empresas (10,6%,  $n=108$ ) y derecho (7,4%,  $n=76$ ) de una universidad pública y cuatro universidades privadas de los estados de San Pablo y Minas Gerais, Brasil. El promedio de edad fue de 23 años y 4 meses ( $DT=5,8$ ), siendo la edad mínima 17 años y la edad máxima 53 años. El sexo femenino representó el 70,5% ( $n=721$ ) de la muestra.

### 1.2. Instrumento

Fue utilizado el texto “*Desentendimento*”, adaptado de Verísimo (1995) y preparado según la técnica de Cloze. El texto posee aproximadamente 250 vocablos y 46 omisiones/espacios. La forma de corrección fue la literal, que es aquella que considera como acierto apenas cuando el espacio es completado con la palabra exacta que fue omitida.

La prueba de Cloze utilizada fue sometida a un análisis morfosintáctico por jueces (dos doctores lingüistas), que clasificaron los ítems de la siguiente forma: adjetivo (8 ítems), adverbio (3 ítems), artículo (3 ítems), conjunción (2 ítems), locución adjetiva (1 ítem), preposición (9 ítems), pronombre (2 ítems), sustantivo (11 ítems) y verbo (7 ítems). La concordancia entre los evaluadores, en la clasificación de los ítems, fue de 100%.

### 1.3. Procedimiento

Los estudiantes que se dispusieron a participar de la investigación, respondieron la prueba de Cloze después de firmar una declaración de consentimiento. Los datos fueron relevados en

las aulas de los estudiantes y cada aplicación demoró aproximadamente 30 minutos.

## 2. RESULTADOS Y DISCUSIÓN

El programa estadístico utilizado para realizar el análisis por el modelo de Rasch fue el Winsteps (Linacre, 2002). Los resultados del análisis de la prueba de Cloze y su ajuste al modelo de Rasch pueden ser observados en la Tabla 1 (página siguiente), que presenta datos de las dificultades de los ítems, los errores patrón, *infits* y *outfits*. Por ser la primera vez que son presentadas las palabras (ítems) que completan los espacios vacíos de la prueba, entre paréntesis se encuentra la palabra en portugués, por haber sido el idioma en el cual fue aplicada.

En la Tabla 1 fue posible observar que la dificultad de los ítems varió de 4,44 a -3,14, correspondiendo respectivamente a la locución adjetiva ‘tradición’ y a la preposición ‘de’. Los ítems ajustados están entre 0,70 y 1,30, siendo que Linacre (2002), considera que 1,5 es el valor límite para admitir un ítem. Los resultados referentes a los ítems pueden ser observados en la Tabla 2 (página siguiente).

En la Tabla 2 se observa que el promedio del *infit*, respecto al ajuste de los ítems, estuvo dentro del patrón esperado, o sea, 1,00. Los valores del *infit* estuvieron entre 1,18 y 0,87, y con eso se puede decir que su variabilidad puede ser considerada entre los niveles esperados, en tanto están dentro del intervalo de 0,70 a 1,30 propuesto por Linacre (2002).

En relación al *outfit*, se constató que su promedio indicó un buen ajuste al modelo. Pese a eso, hubo una variación entre 2,13 y 0,71. Se observó que 7 (15,2%) ítems estaban con valores discrepantes, considerando el parámetro de ajuste (0,70-1,30) de los ítems. La locución adjetiva ‘tradición’, el sustantivo ‘silencio’, los adjetivos ‘polar’ e ‘internacional’, los verbos ‘acompañaba’ y ‘agradecí’ y el adverbio ‘también’ no se mostraron adecuados al modelo. De esa forma, podemos entender que esos ítems fueron respondidos correctamente por personas cuya habilidad no permitiría que los hubiesen acertado.

En un análisis general del instrumento se puede decir que los ítems más difíciles se

Tabla 1.  
Distribución del Ajuste de los Ítems al Modelo de Rasch.

Ítem	Clase	Dificultad	Error	Infit	Outfit
tradição (tradição)	locución adjetiva	4,44	0,38	1,00	1,36
silêncio (silêncio)	sustantivo	4,30	0,36	1,00	1,37
mímica (mímica)	sustantivo	4,30	0,36	1,00	1,09
alejada (afastada)	adjetivo	3,54	0,25	1,00	0,97
brindis (brinde)	sustantivo	3,32	0,22	1,01	1,27
polar (polar)	adjetivo	3,18	0,21	1,02	1,43
internacional (internacional)	adjetivo	3,14	0,20	1,02	2,13
acompañaba (acompanhava)	verbo	2,79	0,17	1,03	1,54
agradecí (agradece)	verbo	2,40	0,15	1,02	1,56
tibio (morno)	adjetivo	2,32	0,14	0,98	0,85
haría (faria)	verbo	2,00	0,12	0,97	0,87
finalmente (finalmente)	adverbio	1,97	0,12	1,01	1,24
también (também)	adverbio	1,83	0,12	1,04	1,37
japonesa (japonesa)	adjetivo	1,79	0,11	0,98	1,07
vamos (vamos)	verbo	1,31	0,10	0,92	0,79
mozo (garçom)	sustantivo	1,20	0,09	0,89	0,71
en la (na)	preposición	1,15	0,09	1,05	1,15
más (mais)	adjetivo	0,90	0,09	1,04	1,07
el (o)	pronombre personal	0,46	0,08	0,98	0,97
y (e)	conjunción	0,35	0,07	1,15	1,28
contentarse (contentar)	verbo	0,13	0,07	0,95	0,91
tenga (tenha)	verbo	0,12	0,07	1,05	1,04
después (depois)	adverbio	-0,49	0,07	0,92	0,99
difícil (difícil)	adjetivo	-0,54	0,07	1,01	1,01
hotel (hotel)	sustantivo	-0,76	0,07	0,97	1,03
cosa (coisa)	sustantivo	-0,81	0,07	1,01	1,05
una (uma)	artículo indefinido	-1,07	0,07	0,93	0,91
el (ele)	pronombre	-1,20	0,07	1,07	1,07
era (era)	verbo	-1,48	0,07	1,18	1,29
la (a)	preposición	-1,57	0,07	1,03	1,02
con (com)	preposición	-1,82	0,07	0,94	0,89
mesa (mesa)	sustantivo	-1,88	0,07	0,87	0,83
con (com)	preposición	-1,92	0,07	1,01	1,05
mozo (garçom)	sustantivo	-1,93	0,07	1,02	1,02
y (e)	conjunción	-2,05	0,08	0,94	0,93
frita (frita)	adjetivo	-2,16	0,08	0,97	0,99
los (os)	artículo	-2,23	0,08	0,98	1,01
restaurante (restaurante)	sustantivo	-2,40	0,08	0,99	0,96
leche (leite)	sustantivo	-2,55	0,09	0,98	0,90
de (de)	preposición	-2,67	0,09	1,00	1,01
la (a)	artículo	-2,69	0,09	1,00	0,93
para (para)	preposición	-2,73	0,09	0,94	0,82
vez (vez)	sustantivo	-2,84	0,09	1,07	1,06
para (para)	preposición	-2,97	0,10	0,91	0,87
del (do)	preposición	-3,03	0,10	0,97	0,91
de (de)	preposición	-3,14	0,10	1,02	0,92

Tabla 2.  
Parámetros de Ajuste de los Ítems.

Parámetros	Infit	Outfit	Error
Promedio	1,00	1,07	0,12
DT	0,06	0,25	0,08
Máximo	1,18	2,13	0,38
Mínimo	0,87	0,71	0,07
1,3 < > 1,5	0 (0,0%)	4 (8,6%)	
1,5 < > 2,0	0 (0,0%)	2 (4,3%)	
< 2,0	0 (0,0%)	1 (2,1%)	

refirieron a la locución adjetiva 'tradição', a los sustantivos 'silêncio', 'mímica' y 'brindis', a los adjetivos 'alejada', 'polar' e 'internacional' y a los verbos 'acompañaba' y 'agradecí'. Los ítems más fáciles fueron las preposiciones 'para', 'del' y 'de'. De esa forma, podemos considerar que esos resultados confirman investigaciones anteriores (Santos et al., 2002; Oliveira &

Santos, 2009) que constataron que había una tendencia, en la prueba de Cloze utilizada en esta investigación (*Desentendimento*), a acertar palabras que se encontraban en las categorías morfosintácticas 'preposiciones' y 'artículos'. Categorías como pronombres y verbos eran más difíciles de acertar.

Con esos resultados fue posible observar que, los ítems que componen la prueba de Cloze estudiada presentan diferentes niveles de dificultad, lo que corrobora los hallazgos de investigaciones anteriores (Abraham & Chapelle, 1992; Oliveira et al., 2009; Oliveira & Santos, 2009; Santos, et al., 2002), o sea, la presencia gradual de dificultad para completar los espacios. Por lo tanto, la prueba estudiada presenta un conjunto de ítems con menor difi-



cultad de respuesta, un conjunto de ítems en los cuales hay una dificultad media de respuesta, y un último conjunto de ítems considerado de difícil respuesta (Abraham & Chapelle, 1992).

En relación a las personas, la mayoría de los valores de *infit* y *outfit* estuvo próxima del promedio adecuado para el modelo, o sea, 1,00. Los datos están presentados en la Tabla 3.

Tabla 3.  
Parámetros de Ajuste de las Personas.

Parámetros	Infit	Outfit	Error
Promedio	1,00	1,03	0,42
DT	0,35	1,23	0,02
Máximo	3,41	9,90	1,84
Mínimo	0,46	0,30	0,41
1,3 < > 1,5	58 (5,6%)	32 (3,1%)	
1,5 < > 2,0	17 (1,6%)	18 (1,7%)	
< 2,0	11 (1,07%)	51 (4,9%)	

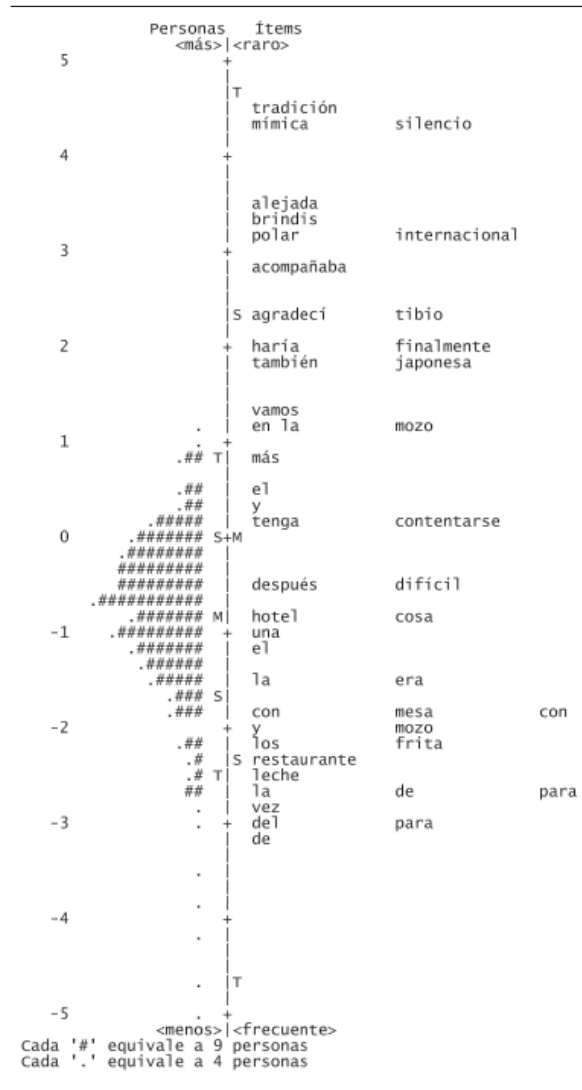
La Tabla 4 presenta el agrupamiento de los ítems y de las personas en la prueba. La letra 'M', al lado de la divisoria de las informaciones de las personas y de los ítems, indica la localización del promedio. Del lado izquierdo se encuentra el promedio de las personas y del lado derecho el de los ítems. Vale recordar que el punto cero en el *Rasch* es el punto de dificultad media, y por lo tanto, los valores negativos son los más fáciles y los positivos y altos son los más difíciles.

Se puede ver que 8,4% ( $n=86$ ) de las personas tuvieron respuestas con valores de *infit* fuera del parámetro esperado. Eso significa que, considerando la habilidad de esas personas, sus respuestas no estuvieron dentro de lo esperado en la prueba de Cloze. En lo que se refiere al *outfit*, apenas 9,8% ( $n=101$ ) de las personas tuvieron respuestas desajustadas. De forma general, los porcentajes de desajuste pueden ser considerados aceptables.

El promedio de las personas fue inferior al promedio de los ítems (Tabla 4). Con eso se puede decir que, de forma general, la prueba de Cloze se presentó como difícil para los universitarios que participaron de esta investigación.

Concluidos los análisis de las propiedades generales de la prueba de Cloze, a continuación será presentado el Análisis del Funcionamiento Diferencial del Ítem (DIF). Es importante destacar que se refiere a un tratamiento estadístico en el cual se estima el DIF, sea positivo o negativo, y cuanto él aumenta en cada ítem. Así,

Tabla 4.  
Mapa de Ítems y de Personas.



por medio de la prueba *t* de *Student* es posible calcular si, la diferencia entre los sexos, puede haber ocurrido o no en razón de una casualidad. Un valor de *t* superior a 2,4 puede ser considerado significativo en función del sexo de los participantes. En la Tabla 5 (página siguiente) son presentados los aumentos de DIF por sexo y los cambios, así como los valores de *t* por ítem. Apenas los ítems con valores superiores a 2,4 (negativo o positivo) presentan dificultad.

Como muestra la Tabla 5, de los 46 ítems analizados apenas cinco diferenciaron el sexo. Los ítems que favorecieron a las mujeres se refirieron al adjetivo 'japonesa', al sustantivo

'mozo' y al adverbio 'después', o sea, esos ítems fueron más fáciles para las mujeres en comparación con los hombres. En el caso de los hombres, los ítems más fáciles fueron el sustantivo 'restaurante' y el verbo 'era'. Como apenas esos ítems discriminaron los sexos, se concluyó que hubo un equilibrio en los sesgos ocurridos para hombre y mujeres. Por lo tanto, no sería necesario sustituir o excluir esos ítems, asimismo, no sería necesario establecer un criterio de corrección diferente para hombres y para mujeres.

De esa forma, se constató que las evidencias de la diferencia por sexo en la comprensión lectora de universitarios, en la cual las mujeres tendrían un desempeño mejor que los hombres, en la prueba de Cloze (Boruchovitch et al., 2007; Oliveira et al., 2003; Oliveira & Santos, 2005, 2006; Silva & Santos, 2004) puede ocurrir en función de otros factores que no se refieren al Cloze favorecer o privilegiar al sexo femenino. Con base en eso se puede decir que la prueba de Cloze estudiada no presentó sesgo de sexo y reveló un equilibrio

Tabla 5.  
Distribución de los Aumentos de DIF por Sexo y Valores de *t* por Ítem.

Sexo	DIF aumentado	Sexo	DIF aumentado	Diferencia	<i>t</i>	Ítem
Masculino	0,12	Femenino	-0,05	0,17	0,86	vez
Masculino	-0,17	Femenino	0,06	-0,24	-1,50	la
Masculino	0,09	Femenino	-0,03	0,12	0,80	hotel
Masculino	-0,68	Femenino	0,42	-1,09	-4,70	japonesa
Masculino	0,02	Femenino	-0,02	0,03	0,22	cosa
Masculino	-0,12	Femenino	0,05	-0,17	-0,83	mozo
Masculino	0,83	Femenino	-0,20	1,03	0,96	silencio
Masculino	0,01	Femenino	0,00	0,02	0,07	de
Masculino	0,53	Femenino	-0,16	0,69	2,18	finalmente
Masculino	0,16	Femenino	-0,07	0,22	1,16	la
Masculino	0,05	Femenino	-0,02	0,07	0,35	leche
Masculino	Puntuación mínima	Femenino	Puntuación mínima	-----	-----	mímica
Masculino	0,03	Femenino	-0,01	0,04	0,23	el
Masculino	0,08	Femenino	-0,03	0,11	0,70	tenga
Masculino	-0,20	Femenino	0,09	-0,30	-1,99	una
Masculino	0,15	Femenino	-0,05	0,20	0,70	haría
Masculino	0,04	Femenino	-0,01	0,05	0,25	de
Masculino	1,26	Femenino	-0,26	1,51	2,04	polar
Masculino	-0,07	Femenino	0,03	-0,10	-0,66	el
Masculino	0,09	Femenino	-0,03	0,12	0,74	y
Masculino	-0,17	Femenino	0,07	-0,24	-1,43	con
Masculino	-0,02	Femenino	0,01	-0,03	-0,14	del
Masculino	0,37	Femenino	-0,12	0,49	1,13	acompañaba
Masculino	-0,08	Femenino	0,02	-0,10	-0,64	mesa
Masculino	-0,19	Femenino	0,08	-0,26	-1,27	para
Masculino	-0,22	Femenino	0,09	-0,31	-1,37	para
Masculino	-0,28	Femenino	-0,09	0,37	1,05	agradecí
Masculino	-0,38	Femenino	0,14	-0,52	-2,99	mozo
Masculino	-0,09	Femenino	0,03	-0,12	-0,72	con
Masculino	-0,27	Femenino	0,11	-0,38	-2,58	después
Masculino	0,90	Femenino	-0,21	1,11	1,47	alejada
Masculino	0,60	Femenino	-0,17	0,76	1,38	internacional
Masculino	0,19	Femenino	-0,08	0,27	1,57	los
Masculino	-0,28	Femenino	0,13	-0,40	-2,08	en la
Masculino	0,71	Femenino	-0,18	0,89	1,42	brindis
Masculino	0,46	Femenino	-0,14	0,60	1,67	tibio
Masculino	0,12	Femenino	-0,04	0,16	0,75	vamos
Masculino	0,70	Femenino	-0,18	0,88	0,81	tradición
Masculino	-0,15	Femenino	0,06	-0,21	-1,14	más
Masculino	0,00	Femenino	0,02	-0,02	-0,12	contentarse
Masculino	0,52	Femenino	-0,16	0,68	2,28	también
Masculino	-0,14	Femenino	0,06	-0,20	-1,13	frita
Masculino	0,38	Femenino	-0,18	0,56	3,20	restaurante
Masculino	0,31	Femenino	-0,14	0,45	2,96	era
Masculino	0,23	Femenino	-0,09	0,31	1,84	y
Masculino	-0,01	Femenino	0,00	-0,01	-0,06	difícil

en la dificultad de los ítems, una vez que los mismos estaban ajustados al modelo. De esa forma, el análisis de los ítems permitió obtener evidencias de validez para el Cloze estudiado, lo que lo torna una medida segura para el diagnóstico de la comprensión lectora de estudiantes universitarios brasileños.

### 3. CONCLUSIONES

Investigar la comprensión lectora y la prueba de Cloze como medida diagnóstica en la enseñanza superior es muy importante, pues es un hecho que, muchos de los estudiantes brasileños -que llegan a ese nivel de la educación formal- no demuestran la comprensión lectora esperada. Probablemente, esto es un punto negativo en la formación de ese futuro profesional, si consideramos que es por medio de la lectura que el universitario conseguirá comprender los contenidos técnicos y científicos. En ese sentido, es posible decir que el diagnóstico de la comprensión lectora, por medio de la prueba de Cloze, podría ofrecer contribuciones para que las universidades busquen medidas correctivas, en aquellos alumnos que muestran dificultades relacionadas a la comprensión lectora, para que puedan recibir el apoyo educacional necesario.

Parece clara la importancia de esta investigación, pues estableció que el resultado ofrecido por la prueba de Cloze investigada, representa la habilidad de comprensión lectora del universitario, demostrando que los espacios/ítems (por su naturaleza heterogéneos) representan el constructo 'comprensión lectora', y que no son susceptibles al sesgo del sexo. Además de eso, debemos pensar que no se puede concluir que, los resultados obtenidos en esta investigación sean generalizados para todos los textos elaborados con base en la técnica de Cloze. Los resultados obtenidos sólo le dan validez al texto usado en esta investigación, siendo que otros textos elaborados con esta técnica también deben ser analizados, en función de sus dificultades.

Los resultados mostraron que los ítems fueron respondidos de acuerdo con el patrón esperado y sin el sesgo del sexo, lo que muestra que este estudio permitió refinar la prueba de Cloze destinada a la evaluación de

universitarios, pues presentó informaciones relevantes sobre sus propiedades psicométricas. También se destaca que existe la necesidad de que nuevas investigaciones sean realizadas.

### 4. REFERENCIAS

- Abraham, R. G., & Chapelle, C. A. (1992). The meaning of Cloze test scores: An item difficulty perspective. *The Modern Language Journal*, 76(4), 468-479.
- Anderson, R. C., & Pearson, D. P. (1984). A Schema-theoretic view of basic processes in reading comprehension. En P. D. Pearson & cols. (Eds.), *Handbook of reading research* (pp. 255-290). New York: Longman.
- Andrade, D.F., Tavares, H. R. & Valle, R.C. (2000). *Teoria de Resposta ao Item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística.
- Araújo, E. A. C., Andrade, D. F., & Bortolotti, S. L. V. (2009). Teoria da Resposta ao Item. *Revista da Escola de Enfermagem*, 43 (N.Especial), 1000-1008.
- Boruchovitch, E., Santos, A. A. A., & Oliveira, K. L. (2007). Análise da fidedignidade entre dois tipos de pontuação do Teste de Cloze. *Psicologia em Pesquisa-UFJF*, 1(1), 41-51.
- Conselho Nacional de Saúde. Resolução 196/96. Diretrizes e Normas Regulamentadoras de Pesquisa. Retirado da Web em 19 de outubro de 2005. <http://www.conselho.saude.gov.br>.
- Cunha, V. L. O., & Capellini, S. A. (2010). Análise psicolinguística e cognitivo-linguística das provas de habilidades metalinguísticas e leitura realizadas em escolares de 2ª a 5ª série. *Revista CEFAC*, 12(5), 772-783.
- Cunha, N. B., & Santos, A. A. A. (2006). Relação entre a compreensão da leitura e a produção escrita em universitários. *Psicologia: Reflexão e Crítica*, 19(2), 237-245.
- Dias, A. S., & Vendramini, C. M. M. (2008). Análise Fatorial com informação completa de uma prova de Compreensão em Leitura em Estatística *Psicologia Escolar e Educacional*, 12(2), 357-367.
- Duke, N.K., & Pearson, P.D. (2002). Effective practices for developing reading comprehension. En A. E. Farstrup & S.J. Samuels (Eds.) *What research has to say about reading instruction* (pp. 205-242). Newark: International Reading Association.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.
- Flippo, R. F. (1998). Points of agreement: a display of professional unity in our field. *The Reading Teacher*, 52, 30-40.
- Gomes, M. A. M., & Boruchovitch, E. (2009). Proficiência em leitura: um panorama da situação. En A. A. A. Santos, E. Boruchovitch & Oliveira, K. L. (Eds.), *Cloze: um instrumento de diagnóstico e intervenção* (pp. 23-46). São Paulo: Casa do Psicólogo.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. En H. Wainer & H. I. Braun (Eds.), *Test Validity* (pp. 339-352). Hillsdale, NJ: Lawrence Erlbaum.



- Hoskins, S. L., Newstead, S. E., & Denni, I. (1997). Degree performance as a function of age, gender, prior qualifications and discipline studied. *Assessment & Evaluation in Higher Education*, 22 (3), 317-328.
- Kintsch, W. (1994). Text comprehension, memory and learning. *American Psychologist*, 49(4), 294-303.
- Linacre, J. M. (2002). What do infit na outfit, nmean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 887.
- Navas, A. L. G. P., Pinto, J. C. B. R., & Dellisa, P. R. R. (2009). Avanços no conhecimento do processamento da fluência em leitura: da palavra ao texto. *Revista da Sociedade Brasileira de Fonoaudiologia*, 14(4), 553-559.
- Nicholson, T. (1999). Reading comprehension processes. En G. B. Thompson & T. Nicholson (Eds.), *Learning to read* (pp. 127-149). Newark: IRA.
- Oliveira, K. L. (2011). Considerações acerca da compreensão em leitura no ensino superior. *Psicologia: Ciência e Profissão*, 31 (4), 690-701.
- Oliveira, K. L., Cantalice, L. M., & Freitas, F. A. (2009). Compreensão em leitura no ensino médio: análise de acertos por item. En A. A. A. Santos, E. Boruchovitch & Oliveira, K. L. (Eds.), *Cloze: um instrumento de diagnóstico e intervenção* (pp. 163-185). São Paulo: Casa do Psicólogo.
- Oliveira, K. L., Boruchovitch, E., & Santos, E. (2009). A técnica de Cloze na avaliação da compreensão em leitura. En A. A. A. Santos, E. Boruchovitch & Oliveira, K. L. (Eds.), *Cloze: um instrumento de diagnóstico e intervenção* (pp. 47-77). São Paulo: Casa do Psicólogo.
- Oliveira, K. L., Santos, A. A. A., & Primi, R. (2003). Estudo das relações entre compreensão em leitura e desempenho em disciplinas na universidade. *Revista Interação em Psicologia*, 7(1), 19-25.
- Oliveira, K. L., & Santos, A. A. A. (2005). Compreensão em leitura e avaliação da aprendizagem em universitários. *Psicologia: Reflexão e Crítica*, 18(1), 118-124.
- Oliveira, K. L., & Santos, A. A. A. (2006). Compreensão de textos e desempenho acadêmico. *Psic-Revista de Psicologia da Vetor Editora*, 7(1), 19-27.
- Oliveira, K. L., & Santos, A. A. A. (2009). Estudo de intervenção para a compreensão em leitura e produção de textos na universidade. *No prelo*.
- Primi, R. (2004). Avanços na interpretação de escalas com a aplicação da Teoria de Resposta ao Item. *Avaliação Psicológica*, 3(1), 53-58.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support student's autonomy during a learning activity. *Journal of Educational Psychology*, 98 (1), 289-218.
- Ruddell, R.B., Ruddell, M.R. & Singer, H. (2001) *Theoretical models and process of reading*. Newark, DE: International Reading Association.
- Santos, A. A. A., Primi, R., Taxa, F. & Vendramini, C. M. M. (2002). O teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, 15(3), 549-560.
- Silva, E. M. T., & Witter, G. P. (2008). Compreensão de texto e desempenho acadêmico em estudantes de psicologia *Estudos de Psicologia (Campinas)*, 25(3), 395-403.
- Silva, E. M. T., & Witter, G. P. (2009). Eficiência de um curso de português no desenvolvimento da leitura e escrita. En A. A. A. Santos, E. Boruchovitch & Oliveira, K. L. (Eds.), *Cloze: um instrumento de diagnóstico e intervenção* (pp. 311-342). São Paulo: Casa do Psicólogo.
- Silva, M. J. M., & Santos, A. A. A. (2004). A avaliação da compreensão em leitura e o desempenho acadêmico de universitários. *Psicologia em Estudo*, 9(3), 459-457.
- Spinillo, A. G., & Mahon, E. R. (2007). Compreensão de texto em crianças: comparações entre diferentes classes de inferência a partir de uma metodologia on-line. *Psicologia: Reflexão e Crítica*, 20(3), 463-471.
- Spira, E. G. Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: the effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, 41(1), 225-234.
- Taylor, W.L. (1953) Cloze procedures: a new tool for measuring reability. *Journalis Quarterly*, 30, 415-433.
- Van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer.
- Venezky, R. L. (1990). Definitions of literacy. En R. L. Venezky, D., A. Wagner & B. S. Ciliberti (Eds.), *Toward defining literacy* (pp. 2-16). Newark: IRA.
- Veríssimo, L. F. (1995). Desentendimento. *Ícaro Revista de bordo da Varig*, XII(136), 11.

Para citar este artículo:

Oliveira, K.L., Rueda, F. J. & Santos, A. A. A. (2012). Ajuste de una prueba de Cloze al modelo de Rasch: calibración, ajuste de los ítems y estudio del DIF. *Ciencias Psicológicas VI* (1): 7-15.

Recibido: 11/2011  
Revisado: 02/2012  
Aceptado: 03/2012