

EL TEST DE CLOZE EN LA ENSEÑANZA PRIMARIA: FUNCIONAMIENTO DIFERENCIAL DEL ÍTEM

CLOZE'S TEST IN PRIMARY EDUCATION: DIFFERENTIAL OPERATION OF ITEMS

Katya Luciane de Oliveira*, Acácia Aparecida Angeli dos Santos**,
Evely Boruchovitch*** y Fabián Javier Marín Rueda****

Universidade Estadual de Londrina, Brasil

Universidade São Francisco, Brasil

Universidade Estadual de Campinas, Brasil

Recibido: 06 de Abril de 2011

Aceptado: 07 de Noviembre de 2011

RESUMEN

Este estudio tuvo como objetivos investigar el ajuste del Test de Cloze al modelo *Rasch* y verificar la presencia de dificultad en la respuesta al ítem en razón del sexo de las personas (DIF). Participaron de la investigación 231 alumnos del 2º al 4º grado de la enseñanza primaria. Se utilizó el test de Cloze, que fue aplicado de forma colectiva. El instrumento se mostró ajustado al modelo *Rasch*. Fue posible observar una buena adecuación y un buen ajuste de los ítems en el *outfit*. En relación al DIF, apenas un ítem indicó diferenciar el sexo. Con base en los datos se concluyó que hubo un equilibrio en los sesgos ocurridos para varones y niñas.

Palabras clave: Test de Cloze, funcionamiento diferencial del ítem, modelo *Rasch*.

ABSTRACT

The objective of this study was to investigate the adjustment of the cloze testing to the Rasch model and verify difficulties in response to the item related to gender (DIF). The sample consisted of 231 students from 2nd to 4th grade Primary Education. Data was collected by means of Cloze test which was collectively administered. The instrument was adjusted to Rasch Models. It was possible to identify in the outfit an adequate functioning of the items. In relation to the DIF, only one item indicated to differentiate by gender. Based on the data, it was concluded that there was a balance in the biases that have occurred for both boys and girls.

Key words: Cloze's Test, Differential Operation of Item, Rasch Models.

Introducción

La lectura es una de las principales formas de obtenerse conocimiento. Su reconocimiento en la vida del hombre es evidente, pues ella proporciona reflexiones y preguntas a respecto de los hechos. La lectura está presente en todos los momentos, siendo indispensable en los días actuales, en especial cuando se considera el potencial de las tecnologías de la información en la vida de las personas (Souza, 2007). En el contexto escolar, el dominio de la lectura

representa el divisor entre el éxito y el fracaso del alumno. Sisto y Martinelli (2006) consideran que el fracaso escolar en esa etapa se refiere, la mayor parte de las veces, a los problemas asociados a la falta de dominio de las habilidades de escrita y lectura.

Datos de la *Organization for Economic Co-Operation and Development* (2001) muestran que la mayor parte de los alumnos de la enseñanza primaria brasileña no tienen un

* oliveira_katya@ig.com.br

buen dominio de la lectura, lo que resulta en una comprensión muy por debajo de lo esperado por los *Parâmetros Curriculares Nacionais* (MEC, 1997), que tienen como propuesta que un alumno de enseñanza primaria debe concluir esta etapa de escolarización con una lectura proficiente. De esa forma, el alumno que concluya esa etapa debe demostrar una comprensión de las ideas más relevantes del texto, comprensión lingüística y comportamiento crítico, de forma que cuando ingrese a la enseñanza secundaria pueda desarrollar una postura crítico-inferencial.

Se puede concluir que la comprensión en lectura no es una tarea simple y que dependería de muchas variables, principalmente de aquellas relacionadas al procesamiento cognitivo que permite la elaboración de esquemas mentales, acceso y recuperación de las informaciones previas y agilidad en la decodificación (Cunha & Santos, 2006; Gilabert, Martínez & Vidal-Abarca, 2005; Ruddell, Ruddell & Singer, 2001; Spira, Bracken & Fischel, 2005). Otros estudiosos argumentan que la comprensión en lectura incluye una articulación equilibrada entre los procesamientos cognitivos y metacognitivos, lo que implica también motivación, planeamiento, monitoreo y regulación del propio comportamiento (Duke & Pearson, 2002; Gonida, Kiosseoglou & Leondari, 2006; Sailor & Abreu, 2005; Sternberg, 2000).

Para Kintsch (1994) la comprensión en lectura incluye la elaboración de analogías por medio de la utilización del raciocinio lógico. Así, una comprensión analítica es el resultado de un buen procesamiento cognitivo. Sobre ese aspecto, Sierra y Carretero (1996) dicen que cuando el lector consigue alcanzar un buen nivel de comprensión la lectura funciona con retroalimentación. Eso significa que la memoria es capaz de acceder a la información ya acumulada, procesarla y reutilizarla. Los estudios de Carpenter, Miyake y Just (1995) demostraron que la comprensión en lectura exige una conexión mental entre las diferentes partes del texto leído, una vez a medida que existe la decodificación y el entendimiento de las sentencias, la comprensión va constituyéndose. En el acto de la lectura, la persona realiza deducciones que llevan a una coherencia global de las partes, resultando en la comprensión.

Hay lectores que utilizan fuentes textuales para obtener una comprensión de la lectura. Esas fuentes pueden ser clasificadas como intratextuales (presentes en el texto), las

intertextuales (compuestas por textos leídos anteriormente) y las extratextuales (externas a la situación de lectura). La comprensión resulta del entendimiento de las relaciones textuales implícitas o explícitas de textos de diversos tipos, tratándose de una habilidad compleja de ser medida (Kopke Filho, 2001).

Frente a la importancia de la lectura y su comprensión, y la necesidad de desarrollar medios que permitan caracterizar el desempeño en lectura de los alumnos de la enseñanza primaria, el test de Cloze está siendo bastante estudiado como una posible técnica diagnóstica. El Cloze fue creado por Taylor (1953) y consiste en un texto de aproximadamente 250 vocablos del cual se omiten todos los quintos vocablos del texto, y en su lugar se coloca una línea de tamaño proporcional a la palabra omitida. Los testados deben completar el espacio con la palabra que juzguen ser más adecuada para el contexto.

En la clasificación de Bormuth (1968) fueron propuestos tres niveles para categorizar el desempeño de los lectores en el Cloze. El primero es identificado como nivel de frustración, y establece que aquellos que obtengan hasta 44% de aciertos no entendieron la información leída. El segundo es el nivel de instrucción, incluyendo puntuaciones que varían de 44,1% a 57%, en el cual el lector muestra una abstracción suficiente para la comprensión, que podrá ser alcanzada apenas con ayuda exterior, que podrán ser pistas adicionales que proporcionen elementos para que a comprensión sea alcanzada. En el último nivel, denominado como independiente, el lector debe obtener un desempeño superior a 57% de los aciertos, lo que demuestra que el lector posee una comprensión crítica, creativa y autónoma del texto.

Con ese sistema de significación se puede decir que las puntuaciones del Cloze han sido interpretadas utilizándose la referencia de criterio, en la cual una puntuación es vista sobre la óptica de alguna otra medida o criterio, como ser los niveles establecidos por Bormuth (1968). Otra forma de interpretar las puntuaciones del Cloze es en relación a los baremos. En ese caso se utilizan los datos del promedio y de la desviación típica. Pese a eso, Page (1975) y Santos, Primi, Taxa y Vendramini (2002) advierten que la comparación de los desempeños en el Cloze debe ser observada con cautela. Ese hecho debe ser considerado, principalmente, cuando el instrumento está fundamentado en la Teoría

Clásica de los Tests, en la cual se considera que el desempeño es producto de puntuaciones calculadas con el presupuesto de que ítems heterogéneos pueden tener el mismo peso. De esa forma, en este estudio existió una motivación para estudiar el test de Cloze con base en la Teoría de Respuesta al Ítem – TRI.

Con la TRI existe la posibilidad de verificar la gradación de la dificultad de los ítems, aunque no se trabaje con muestras representativas. Es posible entender por qué algunos ítems son más difíciles para ser acertados y se puede inferir sobre la habilidad de la persona (Embretson, 1996; Embretson & Reise, 2000). Se trata de un modelo práctico que permite el trabajo con datos dicotomizados, como por ejemplo, los aciertos y los errores del Cloze. Además de eso, no parte del presupuesto que el *rasgo latente* presenta distribución normal. La TRI verificará el *rasgo latente* teniendo como base los ítems. En ese sentido, cada ítem representaría la manifestación del rasgo (Lord, 1980; van der Linden & Hambleton, 1997).

En medio a los modelos propuestos por la TRI se encuentra el de Rasch (1960), que permite investigar si la posibilidad de acierto de un ítem ocurrió solamente en razón de la habilidad de la persona y de la dificultad del ítem. Con base en ello, Sisto (2006a, 2006b) complementa esa información diciendo que un ítem sesgado es aquél que pese a las personas mostrar igualdad de capacidad para responderlo, la perspectiva de éxito en la respuesta correcta es diferente.

En relación a ese aspecto, el interés en utilizar el modelo de Rasch (1960) se basó en el hecho de que algunos estudios (Oliveira, Boruchovitch & Santos, 2007; Oliveira & Santos, 2006; Santos, 1991, 1997) que utilizaron el test de Cloze para identificar la comprensión en lectura de alumnos brasileños mostraron que hay diferencia en la puntuación del Cloze en razón del sexo de los participantes. Pese a que la verificación de tal diferencia no fue el objetivo principal de esas investigaciones, ese análisis fue efectuado e indicó que, de alguna forma, las niñas tuvieron mejores resultados que los varones en los test de Cloze. Esos estudios nos permitieron debatir si la puntuación en el Cloze es sensible a la interferencia de la variable sexo, pues no hay como utilizar un medio de diagnóstico que privilegie uno u otro sexo, pues eso llevaría nuevamente a un sesgo en la medida.

Destaca el estudio de Oliveira, Boruchovitch, Santos y Rueda (2009), que investigó los ítems de un test de Cloze utilizando la TRI y el modelo de Rash. Los autores analizaron un test que fue desarrollado para alumnos de 5° a 8° grado de la enseñanza primaria. Los resultados del análisis mostraron que el test de Cloze presentó una buena adecuación al modelo, así como también que los ítems no privilegiaron a ninguno de los dos sexos. De esa forma, los autores concluyeron que no sería necesario crear baremos específicos de interpretación considerando el sexo.

Por lo tanto, el sexo podría constituirse en un sesgo de medida, siendo que sesgo, a partir de este momento, debe ser considerado como funcionamiento diferencial de los ítems – DIF (Holland & Thayer, 1988), el cual puede caracterizar de forma negativa la calidad del instrumento como un todo, pues el aumento y la disminución de la dificultad del ítem solamente deben referirse a la interacción entre él y la persona. Así, un DIF uniforme se confirma cuando no se observa interacción entre el atributo medido y el hecho de la persona formar parte de algún grupo específico, en el caso de este estudio, sexo. En relación al ítem no uniforme, este ocurre cuando existen probabilidades diferentes de responderlo en los dos grupos (Andriola, 2001; Mellenbergh, 1989; Sisto, 2006a, 2006b, Rueda, 2007).

Con base en las consideraciones presentadas y pensando que a pesar de la eficiencia del test de Cloze como medio diagnóstico de la comprensión en lectura, este instrumento todavía es poco explorado en la enseñanza primaria; esta investigación tuvo como objetivos verificar la aplicación de la TRI en el análisis de un texto preparado de acuerdo con la Técnica de Cloze (test de Cloze), investigando el ajuste del Cloze al modelo *Rasch*, así como el funcionamiento diferencial del ítem (DIF) en relación al sexo de las personas.

Método

Participantes

Participaron 231 alumnos de 2° (28,2%, $n=68$), 3° (34,4%, $n=83$) y 4° (37,3%, $n=90$) grado de la enseñanza primaria, de escuelas públicas estaduais (60,2%; $n=145$) y municipales (39,8%; $n=96$) de los estados de *São Paulo* y *Minas Gerais*. La edad media fue 9 años y 5 meses ($Dt=1,1$). El sexo masculino representó 53,5% ($n=129$) de la muestra y el femenino 46,5% ($n=112$).

* oliveira_katya@ig.com.br

Instrumentos

Fue utilizado un texto de autoría de Santos (2003) preparado según la técnica de *Cloze*. El texto es destinado a niños de 2° a 4° grado de la enseñanza primaria y se intitula «Una venganza infeliz». El texto posee 83 vocablos y 15 omisiones, siendo que Condemarin y Milicic (1988) indican que para niños menores, un texto más reducido puede ser utilizado en lugar de un texto de 250 vocablos. Pese a eso, no establecen un número exacto de vocablos.

Procedimiento

Esta investigación fue presentada y aprobada por el Comité de Ética en Investigación de la Facultad de Ciencias Médicas de la *Universidade Estadual de Campinas*. Todos los procedimientos éticos para la realización de este estudio están basados en las reglamentaciones de la Resolución 196/96 y complementares del Consejo Nacional de Salud de Brasil.

Las aplicaciones fueron realizadas de forma colectiva en las clases de los niños, siendo aplicado el instrumento apenas en aquellos niños que los padres estuvieron de acuerdo con la participación de su hijo, firmando el *contrato de consentimiento libre y esclarecido*. Cada aplicación duró aproximadamente 30 minutos.

Resultados y discusión

Los datos fueron organizados en una planilla y sometidos a estadística inferencial, con la finalidad de alcanzar los objetivos del estudio. El programa estadístico utilizado para realizar el análisis por el modelo *Rasch* fue el *Winsteps*. Los resultados del análisis del test de *Cloze* y su ajuste al modelo *Rasch* pueden ser observados en la tabla 1, que presenta datos de las dificultades de los ítems, errores típicos, *infits* y *outfits*.

Fue posible observar que el promedio del *Infit* en relación al ajuste de los ítems fue adecuado ($P=1$; $DT=0,19$). Se destaca que los ítems ajustados están entre 0,70 y 1,30, siendo que Linacre (2002) considera que 1,5 es el valor límite para aceptarse un ítem. Los resultados referentes a los ítems pueden ser examinados en la tabla 2.

Los resultados obtenidos revelan que los ítems fueron respondidos de acuerdo a lo esperando, es decir, 1,0. Los valores referentes al *infit* estuvieron entre 0,77 y 1,18, por

Tabla 1

Distribución del ajuste de los ítems al modelo Rasch

| Ítem | Dificultad | Error | Infit | Outfit |
|----------|------------|-------|-------|----------------------------|
| casa | 8,40 | 1,83 | | Dificultad máxima esperada |
| nuestras | 3,26 | 0,22 | 1,18 | 3,05 |
| estaban | 2,20 | 0,18 | 1,03 | 0,87 |
| estaba | 1,94 | 0,18 | 1,03 | 0,86 |
| otro | 1,85 | 0,18 | 1,14 | 1,52 |
| quedó | 0,53 | 0,17 | 0,88 | 0,80 |
| muy | 0,39 | 0,17 | 0,85 | 0,89 |
| el | -0,21 | 0,17 | 1,06 | 1,09 |
| ella | -0,33 | 0,17 | 0,82 | 0,61 |
| rasgar | -0,64 | 0,18 | 0,94 | 1,00 |
| su | -1,49 | 0,20 | 0,77 | 0,78 |
| para | -1,57 | 0,20 | 1,10 | 1,18 |
| puerta | -1,77 | 0,20 | 1,14 | 1,03 |
| un | -1,94 | 0,21 | 1,04 | 0,72 |
| las | -2,22 | 0,22 | 1,08 | 0,81 |

Tabla 2

Parámetros de ajuste de los ítems

| Parámetros | Infit | Outfit | Error |
|-------------|-----------|-----------|-------|
| Promedio | 1,0 | 1,09 | 0,19 |
| DT | 0,13 | 0,59 | 0,02 |
| Máximo | 1,18 | 3,05 | 1,83 |
| Mínimo | 0,77 | 0,61 | 0,17 |
| 1,3 < > 1,5 | 0 (0,00%) | 0 (0,00%) | |
| 1,5 < > 2,0 | 0 (0,00%) | 1 (6,6%) | |
| < 2,0 | 0 (0,00%) | 1 (6,6%) | |

lo cual se puede decir que la variabilidad del *infit* puede ser considerada entre niveles esperados, una vez que están entre el intervalo de 0,70-1,30.

En relación al *outfit*, el valor promedio de 1,09 indicó una buena adecuación al modelo. Pese a eso, hubo variación entre 0,61 y 3,05. Se observó que dos ítems (13,3%) estaban con valores discrepantes, al considerarse los parámetros de ajuste de los ítems (0,70-1,30). Los ítems «el» y «nuestras» no se mostraron adecuados al modelo. Se discute que errores en el ítem «el» fueron cometidos por personas cuyas habilidades permitirían su acierto. Ya en el caso del ítem «nuestras», hubo respuestas correctas de personas cuyas habilidades no indicaban la posibilidad de acertarlo.

* oliveira_katya@ig.com.br

Con esos datos se puede suponer que la comprensión en lectura depende de la elaboración de esquemas mentales que permiten el acceso, la recuperación y la utilización de la información (Duke & Pearson, 2002; Ruddell et al., 2001). Se presume también que algunas palabras revelan un procesamiento más complejo y específico (ítems difíciles) y otras requieren una comprensión más global (ítems fáciles). De esa forma, de modo inferencial es posible tener como hipótesis que los ítems del Cloze permitieron que buena parte de los participantes recurriesen a sus conocimientos previos, de forma que pudieron emplearlos en la comprensión del texto. Por lo tanto, los ítems del Cloze, por mostrarse adecuados al modelo, posiblemente llevaron al lector a una coherencia global (Carpenter et al., 1995; Kintsch, 1994), una vez que apenas dos ítems estuvieron fuera del parámetro de ajuste.

En relación a las personas, la mayor parte de los valores de *infit* y *outfit* estuvo próximo del promedio adecuado al modelo, es decir, 1,00. Los datos están presentados en la tabla 3.

Tabla 3

Parámetros de ajuste de las personas

| Parámetros | <i>Infit</i> | <i>Outfit</i> | Error |
|-------------|--------------|---------------|-------|
| Promedio | 0,98 | 1,02 | 0,78 |
| <i>DT</i> | 0,35 | 1,2 | 0,13 |
| Máximo | 2,02 | 9,90 | 1,87 |
| Mínimo | 0,36 | 0,16 | 0,67 |
| 1,3 < > 1,5 | 20 (8,6%) | 11 (4,7%) | |
| 1,5 < > 2,0 | 11 (4,7%) | 14 (6,1%) | |
| < 2,0 | 1 (0,4%) | 22 (9,5%) | |

En el *infit* se constató que una parte de las personas (5,2%; $n=12$) presentó opciones de respuesta desajustadas, pues estuvieron fuera del parámetro. En esa perspectiva, se puede concluir que las respuestas de esas personas no correspondieron a lo esperado en el test de Cloze, teniendo en cuenta sus habilidades para completar el espacio. Para el *outfit*, el análisis mostró que 15,9% ($n=36$) de las personas tuvieron opciones de respuesta desajustadas. De forma general, los porcentajes de desajuste pueden ser considerados aceptables.

Ese resultado puede ser explicado por medio de algunas hipótesis: se piensa que parte significativa de los niños utilizó

las pistas contextuales como, por ejemplo, el hecho de que el tamaño de la línea/espacio en el Cloze fue proporcional al tamaño de la palabra omitida, esto es un recurso intratextual (Kopke Filho, 2001); tampoco se descarta la posibilidad de que los niños hayan utilizado algún recurso extratextual, toda vez que se puede citar el hecho de que algunos de ellos trataron insistentemente de copiar las respuestas de los compañeros. De una forma o de otra, el acierto del ítem no expresaría la habilidad en sí. Esas hipótesis indican que en futuros estudios esas variables en el texto y en la situación de aplicación deben ser controladas con más rigor.

La tabla 4 presenta el agrupamiento de los ítems y de las personas en el test. La letra P, al lado de la divisoria de las informaciones de las personas y de los ítems, indica la localización del promedio. Se destaca que el punto cero en el modelo *Rasch* es el punto de dificultad media, y de esa forma, los valores negativos son los más fáciles y los positivos los más difíciles.

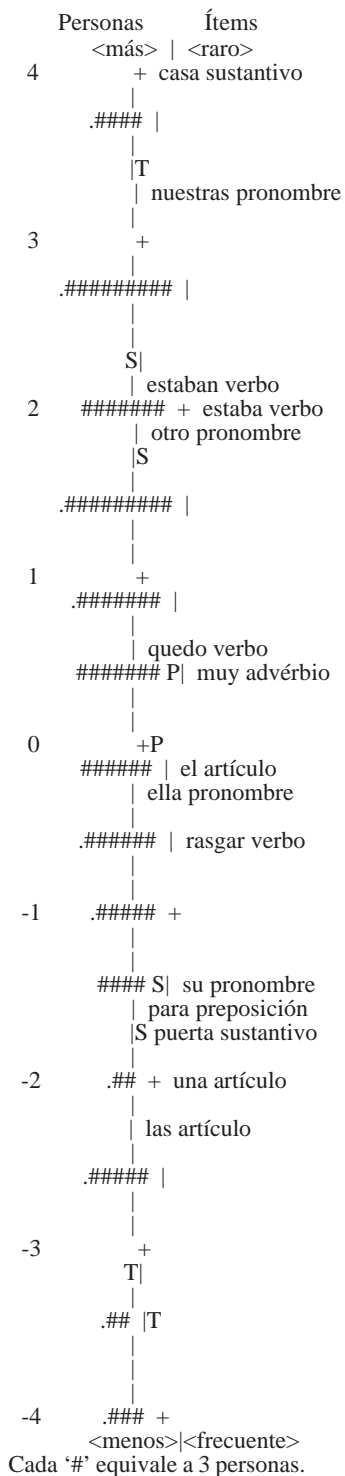
En la tabla 4 es posible observar que el promedio de las personas estuvo próximo al de los ítems. Con esos datos se puede suponer que el promedio de dificultad de los ítems estuvo próximo del promedio de la habilidad de las personas, pese a ser un poco menor, lo que indicaría que el test, de una forma general, se presenta adecuado para ese grupo de niños. El ítem considerado más difícil fue el sustantivo «casa». También existe la necesidad de que en futuros estudios sea realizado un análisis morfosintáctico de los ítems, como una forma de investigar cuáles son las categorías más fáciles o más difíciles.

En relación al análisis del DIF, cabe mencionar que se refiere a un tratamiento en el cual se estima el DIF, sea positivo o negativo, y cuánto él aumenta en cada ítem, así, por medio de la prueba *t* de *Student* es posible calcular si la diferencia entre los sexos puede o no ser en razón del acaso. Sisto (2006a) indica que un valor de *t* mayor que dos (2,0) puede ser considerado significativo en razón del sexo de los participantes. Los resultados pueden ser observados en la tabla 5, donde son presentados los aumentos de DIF por sexo y los cambios y los valores de *t* por ítem.

Puede ser observado que de los quince ítems evaluados, apenas uno indicó diferenciar el sexo. El ítem que favoreció a las niñas se refería al verbo «rasgar», o sea, ese ítem fue más fácil para las niñas que para los varones. Como apenas ese

* oliveira_katya@ig.com.br

Tabla 4
Mapa de ítems y personas



ítem discriminó los sexos, se concluyó que hubo un equilibrio en los sesgos ocurridos para las niñas y para los varones. Así, se consideró que no habría necesidad de sustituir o excluir el ítem, así como tampoco sería necesario un criterio de corrección diferente para varones y niñas, tal como había sido destacado en el estudio de Oliveira et al. (2009), que llegó a la misma conclusión al utilizar un test de Cloze para alumnos de 5° a 8° grado de la enseñanza primaria. De esa forma, los test de Cloze estudiados mostraron medidas consistentes para investigar la comprensión en lectura.

Los datos obtenidos con el análisis del DIF nos llevan a pensar en la hipótesis de aunque las niñas hayan tenido mejores resultados que los varones en investigaciones anteriores (Oliveira et al., 2007; Oliveira & Santos, 2006; Santos, 1991, 1997), no se debe a que los ítems del Cloze privilegien uno de los sexos. Pese a eso, como los estudios citados focalizaron diferentes grados de la educación formal y, por lo tanto, utilizaron otros textos en Cloze, sería ingenuo dejar de mencionar la necesidad de futuras investigaciones que exploren mejor los textos preparados según la técnica de Cloze.

De esa forma, los resultados obtenidos en esta investigación refuerzan la validez del test de Cloze utilizado como medio confiable de medida. Los ítems no mostraron sesgos (Rueda, 2007; Sisto, 2006a), lo que comprueba su eficacia como medio diagnóstico de la comprensión en lectura.

Consideraciones finales

Estudiar la comprensión en lectura es muy importante, considerando que el alumno que presenta dificultades en esa habilidad, muchas veces, fracasa por existir factores que interfieran o no lo ayuden a superar ese problema. Sobre ese aspecto, es posible comentar que el diagnóstico de la comprensión en lectura, por medio del test de Cloze, podría propiciar informaciones más precisas para que el profesor busque alternativas para remediar la dificultad del alumno.

Así, se destaca la relevancia de este estudio, pues debe pensarse en la necesidad de que la medida ofrecida por el Cloze de hecho represente la habilidad de comprensión del alumno. También es necesario mostrar que los espacios/ítems (por su naturaleza heterogénea) representan el constructo «comprensión en lectura» y que ellos no sean susceptibles al sesgo del sexo. De esa forma, con los hallazgos de este

Tabla 5
Distribución de los aumentos de DIF por sexo y valores de *t* por ítem

| Sexo | DIF aumentado | Sexo | DIF aumentado | Cambio | <i>t</i> | Ítem |
|------|---------------|------|---------------|--------|----------|-------------|
| M | 0,04 | F | -0,02 | 0,06 | 0,14 | 1 una |
| M | 0,20 | F | -0,23 | 0,43 | 1,18 | 2 estaban |
| M | -0,42 | F | 0,39 | -0,81 | -0,36 | 4 quedó |
| M | -0,08 | F | 0,06 | -0,14 | -0,41 | 5 ella |
| M | -0,55 | F | 0,37 | -0,92 | -2,44 | 6 rasgar |
| M | 0,21 | F | -0,24 | 0,46 | 1,29 | 7 estaba |
| M | 0,10 | F | -0,06 | 0,16 | 0,40 | 8 para |
| M | 0,31 | F | -0,19 | 0,51 | 1,13 | 9 las |
| M | -0,02 | F | 0,04 | -0,07 | -0,15 | 10 nuestras |
| M | 0,13 | F | -0,12 | 0,26 | 0,78 | 11 muy |
| M | 0,21 | F | -0,14 | 0,35 | 0,89 | 12 su |
| M | -0,13 | F | 0,10 | -0,23 | -0,67 | 13 el |
| M | 0,19 | F | -0,21 | 0,40 | 1,15 | 14 otro |
| M | -0,13 | F | 0,07 | -0,21 | -0,49 | 15 puerta |

estudio se puede decir que el resultado de los participantes en el Cloze se refirió directamente a la dificultad del espacio/ ítem y a sus habilidades, sin ninguna relación con su sexo.

Debe ser mencionado que no se puede concluir que los resultados obtenidos en esta investigación sean generalizados para todo texto preparado con base en la técnica de Cloze. Por lo tanto, los resultados obtenidos ofrecen validez solamente para los textos utilizados, siendo que otros materiales estructurados en Cloze también deben ser analizados en función del grado de dificultad de sus ítems, utilizando la TRI.

Los resultados revelaron que los ítems fueron respondidos de acuerdo con el padrón esperado, lo cual muestra que este estudio permitió un refinamiento del test de Cloze, una vez que ofreció informaciones relevantes sobre sus propiedades psicométricas. Además, sería ingenuo dejar de mencionar que esta investigación tuvo una finalidad exploratoria de los ítems del Cloze, existiendo la necesidad de que nuevas investigaciones sean realizadas con alumnos de otros grados escolares y también de escuelas de naturaleza privada.

Referencias

Andriola, W. (2001). Descripción de los métodos principales para detectar el funcionamiento diferencial de los ítems (DIF). *Psicología: Reflexão e Crítica*, 14(3), 643-652.

- Bormuth, R. (1968). Cloze test readability: Criterion reference scores. *Journal of Educational Measurement*, 5, 189-196.
- Carpenter, P., Miyake, A. & Just, M. (1995). Language comprehension: sentence and discourse processing. *Annual Reviews Psychology*, 46, 91-120.
- Condemarín, M. & Milicic, N. (1988). *Test de Cloze*. Santiago de Chile: Editorial Andrés Bello.
- Conselho Nacional de Saúde do Brasil. Resolução 196/96. Directrices y Normas Regidoras de la Investigación. Retirado de la Web <http://www.conselho.saude.gov.br>
- Cunha, N. & Santos, A. (2006). Relación entre la comprensión de la lectura y la producción escrita en universitarios. *Psicología: Reflexão e Crítica*, 19(2), 237-245.
- Duke, N. & Pearson, P. (2002). Effective practices for developing reading comprehension. En A. Farstrup & S. Samuels (Eds.). *What research has to say about reading instruction* (pp. 205-242). Newark: International Reading Association.
- Embretson, S. (1996). The new rules of measurement. *Psychological Assessment*, 8(4), 341-349.
- Embretson, S. & Reise, S. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum.
- Gilbert, R., Martínez, G. & Vidal-Abarca, E. (2005). Some good text are always better text revision to foster inferences of readers with high and low prior background knowledge. *Learning and Instruction*, 15, 45-68.
- Gonida, E., Kiosseoglou, G. & Leondari, A. (2006). Implicit theories of intelligence, perceived academic competence, and school achievement: testing alternative models. *American Journal of Psychology*, 119(2), 223-238.

* oliveira_katya@ig.com.br

- Holland, P. & Thayer, D. (1988). Differential item performance and the Mantel-Haenszel procedure. En H. Wainer & H. Braum (Eds.), *Test Validity*. Hillsdale, NJ: Lawrence Erlbaum.
- Kintsch, W. (1994). Text comprehension, memory and learning. *American Psychologist*, 49(4), 294-303.
- Kopke, H. (2001). *Estrategias en comprensión de la lectura: conocimiento y uso por profesores de idioma portugués*. Tesis de Doctorado. Universidade de São Paulo, São Paulo-SP.
- Linacre, J. (2002). What do infit na outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 887.
- Lord, F. (1980). *Applications of item response theory to practical testing problems*. Hillsdale: Lawrence Erlbaum.
- Mellenbergh, G. (1989). Item bias and item response theory. *International Journal of Education Research*, 13, 127-143.
- Ministério da Educação e Cultura – MEC (1997). *Parâmetros Curriculares Nacionais*. Retirado de la Web <http://portal.mec.gov.br/seb/arquivos/pdf/livro01.pdf>.
- Oliveira, K., Boruchovitch, E. & Santos, A. (2007). Comprensión textual en alumnos de séptimo y octavo grados de la enseñanza primaria. *Psicología Escolar e Educacional*, 11(1), 41-49.
- Oliveira, K. & Santos, A. (2006). Comprensión de textos y desempeño académico. *Psic*, 7(1), 19-27.
- Oliveira, K., Santos, A., Boruchovitch, E. & Rueda, F. (2009). Comprensión de la lectura: Análisis del Funcionamiento Diferencial de los Ítems de un Test de Cloze. Manuscrito no publicado.
- Organization for Economic Co-Operation and Development. (2001). *Knowledge and skills for life: first results from the PISA 2000*. Paris: OECD.
- Page, W. (1975). The post oral reading Cloze test: new link between oral reading and comprehension. *Journal of Reading Behavior*, 7(4), 385-389.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen, Denmark: Danmarks Paedagogiske Institut.
- Ruddell, R., Ruddell, M. & Singer, H. (2001) *Theoretical models and process of reading*. Newark, DE: International Reading Association.
- Rueda, F. (2007). El funcionamiento diferencial del ítem en el Test Pictórico de Memoria. *Avaliação Psicológica*, 6(2), 229-237.
- Sailor, K. & Abreu, Y. (2005). The nature of memory errors for verbally quantified information. *American Journal of Psychology*, 118(2), 235-250.
- Santos, A. (1991). Desempeño en lectura: un estudio diagnóstico de la comprensión y hábitos de lectura entre universitarios. *Estudios de Psicología*, 8(1), 6-19.
- Santos, A. (1997). Psicopedagogía en la enseñanza terciaria: Evaluación de un programa de remediación en lectura y estudio. *Pro-posições*, 8(1), 27-37.
- Santos, A. (2003). *Una venganza infeliz*. Texto preparado según la técnica de Cloze para niños de los primeros grados de la enseñanza primaria. Manuscrito no publicado. Programa de Pós-graduação Stricto Sensu da Universidade São Francisco. Itatiba.
- Santos, A., Primi, R., Taxa, F. & Vendramini, C. (2002). El test de Cloze en la evaluación de la comprensión en lectura. *Psicologia: Reflexão e Crítica*, 15(3), 549-560.
- Sierra, B. & Carretero, M. (1996). Aprendizaje, memoria y procesamiento de la información: la psicología cognitiva de la instrucción. En C. Coll, J. Palacios & A. Marchesi (Eds.), *Desarrollo psicológico y educación: psicología de la educación* (v. 2, pp. 161-175). Tradução Angélica Mello Alves. Artes Médicas: Porto Alegre.
- Sisto, F. (2006a). Estudio del funcionamiento diferencial de los ítems para evaluar el reconocimiento de palabras. *Avaliação Psicológica*, 5(1), 1-10.
- Sisto, F. (2006b). El funcionamiento diferencial de los ítems. *Psico-USF*, 11(1), 35-43.
- Sisto, F. & Martinelli, S. (2006). El papel de las relaciones sociales en la comprensión del fracaso escolar y de las dificultades de aprendizaje. En F. Sisto & S. Martinelli (Eds.), *La afectividad y dificultades de aprendizaje* (pp. 13-30). Campinas: Vektor Editora.
- Souza, L. (2007). *La importancia de la lectura para la formación de una sociedad consciente*. Retirado de la Web <http://dici.ibict.br/archive/00001095/01/aimportanciadaleitura.pdf>
- Spira, E., Bracken, S. & Fischel, J. (2005). Predicting improvement after first-grade reading difficulties: the effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, 41(1), 225-234.
- Sternberg, R. (2000). *Psicología cognitiva*. Traducción Maria Regina Borges Osório. Porto Alegre: Artes Médicas.
- Taylor, W. (1953). Cloze procedures: a new tool for measuring readability. *Journal Quarterly*, 30, 415-433.
- van der Linden, W. & Hambleton, R. (1997). *Handbook of modern item response theory*. New York: Springer.

* Laboratório de Avaliação e Pesquisa Psicológica/LAPPSIC. Depto. de Psicologia e Psicanálise. Centro de Ciências Biológicas. Universidade Estadual de Londrina. Campus Universitário.