The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test and a reading attitude survey

Maria Cristina Rodrigues Azevedo Joly

Abstract

This study aimed to analyze the efficiency of the Cloze Oriented System (COS) considering its relation with reading attitude and its validity. The Electronic Program Comprehension (EPC) for students from K1 to K4 based on the COS parts of stories from the Brazilian Children's literature was applied in eighteen class-hours. The 40 subjects, ten in each grade, of both genders aged 7 to 11 years, were evaluated before and after the EPC with a Cloze test. They answered the Elementary Reading Attitude Survey too, in a printed protocol adapted to Portuguese, to evaluate the reading attitude of the students and the influence of the EPC in the academic and recreational reading. The results showed that the performance in reading comprehension of all the subjects was higher after taking part in the EPC, and also had significant differences in reading attitude, according to the results of the T of Wilcoxon statistic test. A correlation between comprehension performance before and after the EPC and the reading attitude was verified. There was significant difference for academic reading attitude in the post-test between proficient readers and the one's that has low skill.

Keywords: Cloze Testing; Measurement; Psychometrics.

A validade do Sistema Orientado de Cloze: um estudo correlacional

Resumo

Este estudo analisou a eficiência do Sistema Orientado de Cloze (SOC) considerando sua relação com atitude em leitura e sua validade. O Programa Informatizado de Compreensão (PIC) para alunos de 1º a 4º série do ensino fundamental foi baseado no SOC de trechos de histórias da literatura infantil brasileira e aplicado em 18 horas/aula, duas vezes por semana. Os 40 participantes, 10 para cada série, de ambos os gêneros e com idade entre 7 e 11 anos, foram avaliados antes e depois do PIC por um teste de Cloze. Eles responderam também ao Inventário de Atitudes de Leitura, num protocolo impresso e adaptado para o Português, visando avaliar a atitude de leitura dos alunos e a influência do PIC na leitura acadêmica e recreacional. Os resultados revelaram que o desempenho de todos os participantes em compreensão de leitura foi alto após participarem do PIC e foram encontradas diferenças significativas em relação à atitude de leitura, de acordo com os resultados obtidos pelo teste T de Wilcoxon. Correlação entre o desempenho em compreensão em leitura antes e depois do PIC e atitude de leitura foi verificada. Há diferença significativa da atitude para leitura acadêmica no pós-teste entre os leitores proficientes e os que apresentam pouca habilidade.

Palavras chave: Cloze; Medida; Psicometria.

La validez del Sistema Orientado de Cloze: un estudio correlacional

Resumen

Este estudio analizó la eficiencia del Sistema Orientado de Cloze (SOC) considerando su relación con actitud en lectura y su validez. El Programa Informatizado de Comprensión (PIC) para alumnos del 1er al 4º grado de la enseñanza primaria fue basado en el SOC de pedazos de historias de la literatura infantil brasileña y aplicado en 18 horas/aula, dos veces por semana. Los 40 participantes, 10 de cada grado, de ambos géneros y con edad entre 7 y 11 años, fueron evaluados antes y después del PIC por un test de Cloze. Ellos también respondieron al Inventario de Actitudes de Leitura en un protocolo impreso y adaptado para el portugués, con el objetivo de evaluar la actitud de lectura de los alumnos y la influencia del PIC en la lectura académica y recreacional. Los resultados revelaron que el desempeño en comprensión de lectura de todos los participantes fue alto después de participar del PIC y fueron encontradas diferencias significativas en relación a la actitud de lectura, de acuerdo con los resultados obtenidos por el test T de Wilcoxon. Correlación entre el desempeño en comprensión en lectura antes y después del PIC y actitud de lectura fue verificada. Hay diferencia significativa de la actitud para lectura académica en el pos-test entre los lectores proficientes y los que presentan poca habilidad.

Palabras clave: Cloze; Medida; Psicometría.
Introduction

Literacy and reading comprehension

Literacy is a basic skill and has always been a central concern of schooling. It is necessary to everyone in order to access information, as well as communicating and learning. Especially in the elementary school, the students’ proficiency in reading predicts their success (Braunger & Lewis, 2006). Nonetheless, in Brazil 11.8% of the population aging over 15 years is illiterate. If individuals that have not finished the first four years of elementary school are taken into account, such levels reach 26% (IBGE, 2006).

Moreover, data from the 2003 National Basic Education Evaluation System (Sistema Nacional de Avaliação da Educação Básica – Saeb) revealed that 59% of Brazilian students attending the fourth grade of elementary school display severe reading deficiencies. These students are either illiterate or are still focused on decoding the words rather than understanding their content (INEP, 2003). Only 10% of the students that finish elementary school (8th grade) (INEP, 2003a) and 5% of the students who completed basic education, by finishing the 3rd year of high school, were considered proficient readers (INEP, 2004). Thus, it is necessary to characterize reading performance through strategic evaluations, especially in early schooling stages, in order to develop programs that produce efficient interventions to promote the formation of competent readers.

Beside this, it’s necessary to analyze which type of assessment is effective to know the students’ skill. Johnston (2005) advertises that students need resilient, flexible, self-direct and collaborative literacies to be well-successful in a global world. It’s not only necessary to decode and comprehend the meaning of the language, but also understand how is the best way to interpret and qualify the information from different Medias. At this point of view, despite the complexity of the literacy and the news literacies as constructs, it’s important to consider the assessment relevance and how it will point out the learn directions to teachers (Jonhston & Costello, 2005).

Reading comprehension assessment

Reading and writing are cognitive skills which are part of the Cattell – Horn – Carroll theory. It is comprised of a psychometrically validated model, which integrates ability and academic performance, thus enabling a better comprehension of the learning difficulties in the realm of the psychoeducational variables involved.

As investigated in the present study, reading, is defined as the presence of the necessary abilities to understand written language. Comprehension is related to decoding graphic symbols and assigning meaning to them within a context (sentence, paragraph or text). In order to do that, it is necessary that the linguistic code be recognized, decoded and interpreted (Flanagan et col., 2002).

The international researchers and experts in reading comprehension assessment, according to Braunger e Lewis (2006), agree that the traditional monitoring modes of the reading development during the elementary school are standardized and norm-referenced instruments. Although, theses tests do not have sufficient accuracy to show in detail how is the reader level and his difficulties to become proficient. In addition, their content and format always tend to focus more the right achievement than the substantive learning. Beside this, this kind of evaluation is not common in Brazilian's schools.

The evaluation of reading comprehension, in relation to information interpretation, occurs more frequently through the analysis of reader’s performance in specific tasks in Brazil. Retelling a story that has just been read, identifying the meaning of words, or inferring from context and answering questions are some examples of the tasks most widely used for reading comprehension assessment. Despite the fact that knowledge, application and involvement with information are the main objectives to be achieved by a proficient reader, procedures designed to evaluate knowledge obtained through reading, which is acquired and applied for problem solving, are less common (Paris & Stahal, 2005).

Another procedure for assessing comprehension, called Cloze Technique, was developed by Taylor (1957) to evaluate second language comprehension
and was later adapted by Riley (1986). It requires that the reader establishes relationships among text elements, associations between previous knowledge and printed information, as well as estimate understanding of the contents. Traditional Cloze technique requires the omission of every fifth word in a 250-word text (Taylor, 1957). Progressive Cloze involves the definition of a gradation in the comprehension task sequence, beginning with a stimulus-sentence to get to the text (Riley, 1986). Such an activity is oriented by the teacher to guarantee that the students master the procedure and apply it to routine reading situations. Another variation of Cloze is related to the quantity of words omitted of a particular semantic or syntactic category of information [10, 11]. The task is to fill out the blanks in such a way that the meaning of the sentence or text is preserved. It is considered ability in the CHC theory (Flanagan e col., 2002).

The Cloze Oriented System (COC) is used in evaluation instruments comprised by literary texts (Joly, 2006; Joly & Lomônaco, 2003; Joly & Nicolau, 2005) and deals with text organization. It could be specified from a specific set of criteria related to the number of words, word deletions, size of blanks, and options for response. It aims at determining different difficulty levels.

Studies with Cloze

It must be pointed out that Cloze technique is efficient for developing and implementing reading comprehension. Nonetheless, few studies were made available in the past decade, in Brazil with 1st to 4th grade elementary school students (Joly & Marini, 2006).

There are some Brazilian studies with intervention procedures, such as, the one from Joly (2000), focused on the analysis of the effect of a reading comprehension program, Joly and Lomônaco (2003) which compared the effects of the media used (printed or computerized) in a reading program, as well as Santos (2004), which aimed at analyzing the psychometric characteristics of the test used for the investigation. There are other studies, which are aimed at designing reading evaluation studies using Traditional Cloze (Taylor, 1957).

One of the psychometric studies of a comprehension test (Cloze – MAR), using Traditional Cloze, applied to text adapted from children’s literature, was conducted, by Joly and Nicolau, with 511 students, between 9 and 14 years of age (M=9.80; SD=8.40), of which 53 % were male, attending K4, in public and private schools of the interior of São Paulo state/Brazil. Evidence of the validity of the construct, in relation to age and criterion by extreme groups, was found for the reading comprehension in the Cloze used. As for its reliability, Cronbach’s coefficient test indicated a reliability of 0,95 (Joly & Nicolau, 2005).

This test (Cloze – MAR) was also analyzed through Item Response Theory (IRT) with two parameters (Joly, 2006). The respondents were 522 K4 students aged between 9 and 14 years (M=9.82; SD=0.87), 53.4% male. They were attending public (57%) as well as private schools (43%) in the interior of São Paulo state. The Kuder-Richardson test displayed a reliability of 0.95. The average difficulty found for the items was 0.81 (SD=1.16), and was considered adequate for the two-parameter model. Of the total existing 59 items, 25 were considered difficult, because they displayed values above 1.50, with a critical value of 2.95. Item discrimination index was 1.04 (SD=0.25), which was excellent, for it is much superior to the critical value of 0.30. This thus indicates that difficult items require superior abilities from respondents and vice-versa (Embretson & Reise, 2000). It is worth pointing out that the high reliability of the test remained unaltered (0,95/0,94) independently of the statistical test used, as shown by studies presented earlier.

A test (Cloze – MAL) using COS – 1 x 5 (every fifth word deleted) by box (word list) and by option (multiple choice with three options) applied to an adapted children’s literary text, to evaluate the comprehension performance of K2 and K4. It was analyzed through the Classical Statistics Theory (Caparrotti, 2005) and IRT Two parameter model (Embretson & Reise, 2000; Joly, 2006).

Caparrotti (20005) found convergent-discriminant validity with the Peabody Vocabulary Test in an investigation with 724 students with average age of
The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test and...

Maria Cristina Rodrigues Azevedo Joly

52

9.65 (DP=0.81) in K3 and K4, attending both public and private schools, using COS -1x5 by box. The test discriminated proficient and non-proficient readers according to variables of gender, grade, age and type of school according to Peabody. Reliability was 0.94 using Cronbach’s coefficient test.

A test per box, in which 1220 participants were instructed to choose the word to complete the blank in the text, from a list that contained all the words which had been omitted, without distracters, was analyzed by using IRT. Participants average age was 9.49 (SD=0.95). There were 51% males K2 (8.4%), K3 (47.5%) and K4 (44.1%) attending public (77%) and private (23%) schools. The average difficulty of the items was 0.09 (SD=0.62), which is adequate for this model. Test items are easy, but have a very good average discrimination index (M=1.06; SD=0.40) according Embretson and Reise (2000). The reliability identified was also very good (KR-21 = 0.94).

The study of the option test using IRT was conducted with 275 K2 students, with average age of 8.33 (SD=0.88), 50.2% of which were female, and 84% attended public schools. A reliability of 0.92 was detected by Kuder-Richardson test. The average difficulty found for the items was 0.28 (SD=0.62), which is satisfactory for the model. Test items can be considered of average difficulty and high level of discrimination (M=1.08; SD=0.01), according Embretson and Reise (2000).

Comparing the studies, the precision was equivalent in the investigations using tests by box (0.94) and there was no significant decrease (0.2) by option (Joly, 2006; Caparrotti, 2005). The results have shown that the box test is easier to be answered than the multiple choice test. The different difficulty levels of the two tests, although the contents are the same, due to the variation of response option, revealed that it is possible to use COS as a system for organizing the tests, by defining difficulty levels.

Thus, the problems for understanding the text will be linked to the items omitted and to the type of answer required, as it is the case in the Degrees of Reading Power Test, text difficulty is closely linked to its readability (Koslin, Zeno & Koslin, 1987). Is must be pointed out that COS was investigated solely as applied to an evaluation instrument, and was found to be viable.

Another important aspect, which influences reading proficiency, is reading attitude. This is defined as a set of feelings in relation to reading that determines student’s adhesion or withdrawal when faced with reading situations. Such attitude is derived from reading experiences the reader has had, either recreational or academic; and is related, on one hand, to social norms and subjective beliefs about what it means to be a good reader, as well as motivation, and on the other hand, to cognitive and metacognitive abilities to understand written language (Mathewson, 2001).

There are psychological models, which have been developed about reading attitude acquisition (McKenna, 1994). It’s necessary to define the attitude object of reading which will be evaluate in order to analyze the results to the development of proficient readers and then identify the contexts and materials that influence the performance (Ajzen, 1989).

Researches and surveys, in which the Elementary Reading Attitude Survey was used with elementary school students, showed that children’s attitudes about reading exhibited a consistent decline in function of age. The young readers have more positive reading attitudes than the oldest ones. Besides this, the girls expressed more positive attitudes towards recreational reading and greater stability than boys (Kush & Watkins, 1996; McKenna, Kear & Ellsworth, 1995; Sainsbury & Schagen, 2004).

Morgan e Douglas (2007), after analyze 15 studies, identified that reading skills are correlate with motivation and reading attitude. In one hand, young children who enjoy reading do it more often and they tend to become skilled at it. In the other hand, poor readers often display low motivation to read. According Guthrie e cols. (2007), the motivation to read, considering interest, perceived control, collaboration, involvement, and efficacy could predict the reading comprehension growth.

In addition, characterizing students’ reading attitude and assessing the possible relationships and influences in comprehension performance is always a
contemporary field of study for literacy research. In order to attempt this tendency and minimize the reading tests gap to assess the students’ achievement, this study evaluated the validity of the Cloze Oriented System (COS), when applied to a reading intervention program, as a reading comprehension assessment procedure. It was aimed by comparison of pre and post test situation, differentiation by extreme groups and correlation with reading attitude.

**Method**

**Participants**

40 students of both genres (57.5% female) attending K1 to K4. It was 10 students per grade, participated in the study. The age varied between 7 and 11 years (M=9.05; SD=1.20).

**Instruments**

*Electronic Program of Comprehension (EPC)* (Joly, 1999)

Its goal is to develop reading comprehension ability. It was designed for K1 to K4 students. Its organization is based on COS applied to 16 adapted excerpts of Brazilian children’s literature. The texts varied in size (150 to 300 words), deletion criterion (every 10, 8, 7 or 5 words) and response option (2 or 3 options, list of omitted words, first letter of the word deleted, number of letters of the word deleted or no option), and were divided in four steps with an increasing level of difficulty. The activities of the program can be done by computer, individually or in pairs, thus respecting each student’s rhythm. Evaluation is done by summing the correct responses. The participant must obtain 70% of correct responses to go to the next story and step.

*Reading comprehension test* (Joly, 1999).

A 300-word text, in Traditional Cloze format, with omission of one in every five words was selected. The blanks were to be filled out with the words that the participant considered as being the best to complete the meaning of the text, without options. The answers considered correct were the ones that had words identical to the original text. One point was awarded for each correct response.

*Elementary Reading Attitude Survey* (McKenna & Kear, 2000).

This indicates the attitude towards both recreational and academic reading of elementary school students. The survey is comprised of 20 items in the form of questions that always begin as “How do you feel...” (E.g. How do you feel when you get a book as a present) in a four-point Likert scale which varies from “Very happy” (3 points) to “Sad” (zero point), that was translated to Portuguese. The first ten questions refer to recreational reading and the remaining questions to academic reading. Score is obtained by the total sum of the frequencies per item and factor, with a maximum possible score of 80 points. The higher the score, the more favorable the attitude of the respondent towards reading will be.

**Procedure**

The participants were initially evaluated through the Reading Comprehension Test and by the Elementary Reading Attitude Survey. The EPC was conducted in a maximum of 16 sessions, with duration of 50 minutes, twice a week. The participants did the activities independently. During the conduction of the activities of the program a tutor provided any necessary orientation, about equipments or technical issues. After the completion of the program, a new reading comprehension evaluation was conducted.

**Results and Discussion**

The analysis of the reading comprehension performance revealed progress in all grades, which corroborates previous studies, using Cloze technique in non-computerized comprehension programs (Joly & Lomônaco, 2003), with students beginning schooling (Joly & Lomônaco, 2003, p.16-17). The worst pre-test performance was in K1 and the best post-test performance was in K3 (Figure 1).
Wilcoxon T-test indicated a significant difference between pre and post test for all grades (Table 1). A correlation between comprehension performance before and after EPC as detected \((r=0,26; p<0,05)\). This probably occurs because the evolution in acquisition of the skills was extensive for all grades. Caparrotti (2005) also detected an age difference with COS for K4 in her study.

![Reading performance to pre and posttest by grade](image)

**Table 1: Wilcoxon T-test analysis of reading comprehension performance in pre and post test by grade (N=10)**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Pre test</th>
<th>Post test</th>
<th>Wilcoxon T-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>K1</td>
<td>0,20</td>
<td>0,62</td>
<td>39,00</td>
</tr>
<tr>
<td>K2</td>
<td>6,50</td>
<td>8,63</td>
<td>42,60</td>
</tr>
<tr>
<td>K3</td>
<td>1,20</td>
<td>3,79</td>
<td>48,90</td>
</tr>
<tr>
<td>K4</td>
<td>8,40</td>
<td>2,41</td>
<td>45,60</td>
</tr>
</tbody>
</table>

**significant to \(p<0,05\)**

ANCOVA applied to the comprehension score (pre-test) in relation to variables of age, gender and grade attended, showed an effect of grade upon comprehension performance \((F[3]=6,00; p<0,003)\), with an interaction of gender and age \((F[3]=4,32; p<0,01)\). The analysis of post-test scores did not detect any effect of the variables on performance.

The Elementary Reading Attitude Survey results showed an increase in the total score (Figure 2), without any differentiation by grade for pre-test \((M=57,50; SD=13,28)\), and for post-test \((M=65,83; SD=10,99)\).

Such difference was significant for K1 \((Z=-2,08;p<0,005)\), K2 \((Z=-2,65;p<0,008)\) and K3 \((Z=-0,71;p<0,008)\).
2.90; \( p < 0.037 \). No difference was identified for K4 (Z = -0.59; \( p < 0.55 \)). The reliability index was 95%.

ANCOVA indicated an effect of grade on reading attitude both for pre (\( F[3]=3.36; \ p < 0.035 \)) and post-test (\( F[3]=3.27; \ p < 0.038 \)).

There were not only correlation between reading comprehension and reading attitude to pre and post-test, but also to recreational and academic reading. It was very high significant correlation between reading comprehension and reading attitude and academic

<table>
<thead>
<tr>
<th>Reading comprehension</th>
<th>Index</th>
<th>Reading attitude</th>
<th>Recreacional reading</th>
<th>Academic reading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>pre-test</td>
<td>pos-test</td>
<td>pre-test</td>
</tr>
<tr>
<td>Pre-test</td>
<td>( r )</td>
<td>0.384(*)</td>
<td>-</td>
<td>0.351(*)</td>
</tr>
<tr>
<td></td>
<td>( p )</td>
<td>0.007</td>
<td>-</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>( N )</td>
<td>40</td>
<td>-</td>
<td>40</td>
</tr>
<tr>
<td>Post-test</td>
<td>( r )</td>
<td>-</td>
<td>0.416(*)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( p )</td>
<td>-</td>
<td>0.004</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>( N )</td>
<td>-</td>
<td>40</td>
<td>-</td>
</tr>
</tbody>
</table>

\* \( p < 0.05 \)  \** \( p < 0.001 \)
reading after the EPC; to recreational reading the correlation is almost the same (Table 2).

Only in relation to K2 a correlation between performance after EPC and academic reading attitude prior to EPC ($r=0.64; p<0.05$) was detected. The participants were sorted according to best performance (average total score of at least 75% of the maximum score obtained) and worst (average total score lower than 25% of the maximum score obtained) so that an analysis of reading attitude could be conducted.

A high level of correlation was identified among participants with poor comprehension performance prior to EPC and academic reading attitude after this intervention program ($r=0.84; p<0.05$). There was a statistically significant difference in relation to academic reading attitude post-test by extreme groups of comprehension performance to K2 participants ($t[4]=-2.95; p<0.04$). It was identified that reading attitude predicted (pre-test) the achievement in reading comprehension ($t=2.64; p=0.012$) observed in the post-test using the linear regression test with the backward method.

K2 was the grade, which displayed the most significant relationships between comprehension performance and reading attitude in relation to school. We can assume that such a relation is due to reading acquisition, if we consider the integration between decoding and comprehension (Flanagan et al., 2002), that occurs in this grade in Brazil (MEC, 1996), which enables the reader to get involved in this activity. More over, it is very important to characterize readers and verify the viability of implementing reading comprehension programs for reader formation in this initial schooling phase (Paris & Stahal, 2005).

In relation to the objectives proposed, the efficiency and validity of Cloze Oriented System (COS) applied to a reading intervention program was verified by means of comparing pre and post test status, which revealed significant gain for the participants in all grades. Evidence of validity differentiation in reading attitude of the participants, by extreme groups of comprehension, and convergent-discriminant validity, by correlation between the EPC and the Elementary Reading Attitude Survey, was found.

### Final Remarks

COS has shown that it is an efficient Cloze technique both for application in evaluation and in computerized intervention programs for developing reading comprehension. Despite the small number of participants, it is possible to obtain significantly favorable results for this text organization system. Further studies with IRT will be necessary in order to fit the model and estimate item difficulty and student skill in the EPC. The relationships among variables such as memory, attention, processing and reaction time, are also relevant to be investigated, especially for a computerized program.

### References


About the author

Maria Cristina Rodrigues Azevedo Joly (cristina@trtec.com.br): Universidade São Francisco Rua Alexandre Rodrigues Barbosa, 45 - 13251-900 Itatiba/ São Paulo, Brasil.