How EFL Learners Fill in the Blanks of an X-test: Think-Aloud Protocol

Hamid Ashraf1, Mona Tabatabaee Yazdi1 & Aynaz Samir1
1English Department, Torbat-e-Heydarieh Branch, Islamic Azad University, Torbat-e-Heydarieh, Iran
Correspondence: Mona Tabatabaee Yazdi, English Department, Torbat-e-Heydarieh Branch, Islamic Azad University, Torbat-e-Heydarieh, Iran. E-mail: Tabatabaee.mona@gmail.com

Received: July 27, 2016   Accepted: August 20, 2016   Online Published: January 12, 2017
doi:10.5539/ijel.v7n1p178       URL: http://dx.doi.org/10.5539/ijel.v7n1p178

Abstract
Since SLA literature remains researchers unaware of the mental processes involved in the X-Test taking (in contrast to C-Test which there are plenty of available related studies), this article aims at exploring cognitive strategies that EFL learners may use while answering an English X-test, which like the C-Test has been modified, adapted and used in many research papers. To this aim, thirty EFL respondents from Mashhad, Iran, were randomly asked to answer a reliable and valid X-test. All of them participated in introspective methods of think-aloud and retrospective interviews during and after the test administration. To analyze the data only the exact word scoring procedure was employed. The results showed participants used various cognitive strategies in taking the X-Test. It was also revealed that respondents experienced more strategies when filling out an X-Test comparing to related literature of C-test, which could be an indicator of the importance job of cognition in X-Test taking. It is hoped that the article can shed light on the underlying cognitive strategies that English language learners‘ use, and provide a chance for educators who want to better understand the learners‘ cognitive processes in order to assist them identify problems and improve their English instruction.

Keywords: X-test, modified C-Test, EFL learners, think aloud protocol

1. Introduction
In the area of language and language testing one of the primary questions for the psycholinguist is how far the test prompts authentic language behavior from the participants. Undeniably, one of the key criticisms of multiple-choice assessments is that putting crosses in boxes does not have any correlates with authentic language use. Besides, one of the concerns, most frequently focused on C-Tests and modified C-Test known as X-Test is whether the behavior extracted from the texts can be considered as a sign of the respondents‘ general language proficiency.

What goes on when a person takes an X-Test? To answer this question a deep investigation of test-taking processes is required. Grotjahn (1986) proposes three possible approaches to this aim: statistical item analysis, text linguistic item analysis and analysis of individual performance. Concerning the third approach, this paper provides evidence derived from an in depth cognitive analysis of participants‘ responses and their mental strategies to be used while engaging in filling out an X-Test.

2. Literature Review
2.1 X-Test and Think Aloud Protocol
Teaching and testing reading comprehension are crucial areas in the field of language teaching and testing both in EFL and ESL contexts. Since different language professionals do not have the same ideas of what reading is and how it can be tested, assessing of reading comprehension is done by different means to evaluate different kinds of abilities. One of such means is known as C-test.

The Modified C-Test (the MC-Test), also known as the X-Test and left-hand deletion was initiated by Boonsathorn (1987, cited in https://www.ukessays.com). The original Modified C-Test is a test that the first half of every second word is deleted and the students are asked to fill in all the blanked parts. In the X-Test, if the deleted word contains an even number of letters, the first half of this word must be deleted. For a word with an odd number of letters, its larger part must be removed.
The use of modified versions of C-tests is not something new and examples of applying such forms are cited in the literature. For instance, the problem of poor discrimination of C-tests was asserted by Cleary (1988) through employing a C-test in which left rather than the usual right grammatically marked items were removed. The results of the study revealed that the discrimination of the C-test could be improved by left-hand deletion. In another study, Heidari (1999) made a comparison between a left-hand C-test with an original one and showed that the C-test with left hand deletion enriched the discrimination power of the C-test.

Over the years, numerous evidence of validity including invariance of item parameters (Baghaei, 2010), independence of items (Eckes & Baghaei, 2015), fit to latent trait models (Baghaei & Grotjahn, 2014a, 2014b; Baghaei, 2008; Baghaei, 2014c; Eckes & Grotjahn, 2006), and correlational evidence have been accumulated for C-Test (see Sigott, 2004). In view of that some other research findings such as Prapphal (1996) state that the X-Test had high reliability and validity in assessing grammatical competence, and the language proficiency of English for non-native-speaking test-takers (Boonsathorn, 1987, cited in https://www.ukessays.com). Prapphal (1996), also created two X-Tests by using texts of General English and Academic English to determine whether the X-Tests in the study could better assess lexical competence or the grammatical competence. The results showed that an X-Test constructed from General English or Academic English is extremely reliable and valid in measuring the grammatical competence.

Considering studies reported above, one can conclude that all of them have been concerned with quantitative analysis of the data gathered from using different forms of C-tests and few studies have been assumed to observe such data qualitatively. Meanwhile, there are many studies in the literature that attempts to measure learning strategies in different contexts with various data gathering procedures (Schellings, 2011; Scott, 2008). One of these strategies which has been frequently applied and specifically related to learning from text is think aloud (Caldwell & Leslie, 2010 Fox, 2009; Greene, Robertson, & Croker Costa, 2011).

Think-aloud protocol is used to examine cognitive psychology thinking (Crutcher, 1994), cognitive science (Simon & Kaplan, 1989), and analyzing behavior (Austin & Delaney, 1998), educational psychology (Pressley & Afflerbach, 1995; Renkl, 1997) and design research (Gero & McNeill, 1998) by asking participants to verbalize their thoughts when reading, or answering to questions proposed by teachers. Learners might verbalize explanation, questions, try to predict, or draw conclusions.

The protocol and familiarity of learners with the habit of thinking aloud boost classroom discourse and give learners an opportunity to learn how to learn and provides teachers with the strengths and weakness of the learners. However, one of the most important benefits of think-aloud protocol is to help students expand the ability to monitor their reading comprehension and facilitate their understanding.

Baumann, Jones, & Seifert-Kessell (1993) suggested that think-aloud strategy is typically used to make predictions, create images, review and connect information in text with previous knowledge. This metacognitive awareness which aims at helping learners to be able to think about their own thought, monitor their level of comprehension and modify their strategies for future success is an important component in learning (Oster, 2001).

Therefore, the present study was conducted to fill such an obvious gap in the literature by exploring the way the participants reacted to an X-Test through conducting think-aloud protocol. Consequently, the following major research question was formulated:

What cognitive strategies may respondents use in answering an X-test?

3. Method

3.1 Participants and Setting

Since the purpose of each design is to try to avoid possible errors, in order to share findings with others, therefore this study adopted a qualitative method design using Think-Aloud protocol. A total number of 30 Iranian university learners participated in this study. They were both males (60%) and females (40%) and aged between 20 and 30 years old. All of these participants were Persian native speakers. Participants’ homogeneity in terms of language proficiency was recognized through their final scores of their General English course. Selection was done randomly from students who were studying for their bachelor’s degree at Azad university of Mashhad. Their field of study was Teaching English as a Foreign Language (TEFL). Consequently, the sample seems to have homogeneity regarding age, English language background, L1 background and educational level.

3.2 Materials

Data gathering started in February 2016 and lasted about four months using an X-Test which was extracted from
http://www.ukessays.com. The reliability of the tests was also estimated by Boonsathorn (1990) and Wonghiransombat (1998) and the result asserted a high reliability and validity of the tests. Furthermore, participants were asked to notify their demographic information including age, and gender.

3.2.1 Procedure

To pursue the goals of the current study, thirty participants who were all Iranian ELT university students from Mashhad, Iran, were selected randomly. They were from both genders and different ages.

For collecting the data, an X-Test is administrated in the form of paper. Collecting data started in February 2016 and because data were collected individually from each subject it lasted for about 4 months.

Researchers presented the test to participants and asked them to think aloud and verbalize whatever comes into their mind as they fill the blanks. This gives researchers insight into the participant’s cognitive processes in addition to their final product.

To analyze the data, subjects’ expressions were recorded using a tape recorder and then transcribed. The assigned time for completing the test was 15 minutes. If participants forgot to think aloud, researchers provided them with some open-ended questions such as “What are you thinking now?” Moreover, to illuminate any misconception that might have happened during the think-aloud researchers informally interviewed students by asking some questions such as “Can you explain what you meant?”

As Baumann et al. (1993) stated participants usually verbalized some sentences orally, such as: “So far, I’ve understood that...”; “This made me think of...”; “I think.... will happen next.”; “I reread that part because...”; “I was confused by...”; “I think the most important part was...”; “That is interesting because...”. By these verbalization participants typically use the following strategies: Questioning, Predicting, Clarifying, Making Connections, Re-reading/Fix-ups, Visualizing, Summarizing, and Commenting.

Finally, a rubric was used as an instrument to study each student’s think-aloud. The researchers then tried to conclude from the protocols the mental strategies which were applied by the test takers for the individual test item. Furthermore, gathering data was summarized by the use of SPSS software (Version 16). Then, descriptive statistics were reported as the analysis of the data.

4. Analysis of Data and Results

To discover any significant use of different cognitive processes applied by participants for filling the gaps of the X-test, a rich source of information, through participants’ verbalization (think aloud protocol), was gathered by the researchers.

Data transcriptions were coded into several themes of cognitive strategies such as “Making prediction”, “Summarizing”, “Making question”, “Using connection”, “Using fix-ups”, “Identifying a problem”, and “Reflecting”.

Descriptive analysis of the results revealed that the above mentioned seven cognitive strategies were frequently used by the participants, though their usages were not equal statistically (Table 1).

<table>
<thead>
<tr>
<th>Strategies</th>
<th>Making prediction</th>
<th>Summarizing</th>
<th>Making Q</th>
<th>Using connection</th>
<th>Reflecting</th>
<th>Used fix-ups</th>
<th>Identify a problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-Test</td>
<td>61%</td>
<td>17%</td>
<td>7.5%</td>
<td>6%</td>
<td>3.5%</td>
<td>3%</td>
<td>2%</td>
</tr>
</tbody>
</table>

According to the results, making prediction was revealed to be the most frequently used strategy for answering the test. 61% of the correct answers were filled by using “Making prediction” that means participants used clues to make guesses about what they were reading. In case of “the bottom -e bottom -f the -ip”, a test taker stated, “The bottom of the ship...the bottom... I think that the words “the” is correct. I chose the because of bottom, I think the sentence lacks article”

The second frequently applied approach was related to “Summarizing” or taking the most important information of the text by participants and putting it in their own words. Faced with a sentence like: “they -iment with -rent kinds -f Thai -od and -find that -t tastes -ious” One of the subjects verbalized as follows: “-rent kinds of Thai -od”...; the paragraph talks about Thailand and focuses on different parts of Thailand then it started to talk about its hotels when I saw the word Thai and also the tastes, I realized that it focuses on food so...” they experiment with different kind of Thai food and it tastes delicious”.

Next, 7.5% of the X-test was truly answered by the participants through “Making question” strategy in which
students asked questions about what they did not know and tried to find the answers. For instance, when one student faced with “the -hes are -an”, he instantly said “the question I have is what is “the -hes are -an”? What words can I make with these letters?”

The forth commonly applied cognitive method (6%), which was very close to “Making question” in frequency of usage was “Using connection” in which participants thought about how things they were reading, connected to their lives, the world around them, or what they have read. For instance, in case of “-en we -ch a -ip”, a subject stated, “The first word is then or when because it is similar to the words which ending in en”.

“Reflecting” was ranked by the respondents in the fifth position with 3.5% of usage. For instance, when faced with “disappears -st, and -en the -ip”, the subject said, “I wrote first and then because of the events that occurs in sequence”.

3% of the text’s gaps were filled by using “Fix-ups” that is rereading the text and co-test. For example, in case of, “-at is -ing the -ip”, a student stated, “-at is -ing… I’ll reread the first and the second word”.

Finally, the last frequently used strategy (2%) was identified to be “Identifying a problem”. Participants used this strategy by clarifying and verbalizing some questions such as “I am confused about…, … I didn’t expect…, I am not sure of….. For instance in “-ls out -o sea”, a test taker said, “I’m not sure what exactly “-ls” is, but I know it is something about the ship and the sea.

To summarize, the study discovered that in filling out the gaps of an X-Test, seven different cognitive strategies were used by the participant. Although all the strategies were used by each individual participant, the strategies’ frequency of occurrence was not similar. The most frequently used strategy was “Making prediction” by 61 percent of usage while “Identifying a problem” was shown to be the least frequently applied strategy by 2 percent of usage.

5. Discussion and Conclusion

One of the crucial tests which has been used to measure reading comprehension is C-Test, which was introduced in 1981 (cited in https://www.ukessays.com). C-Test has been modified and revised into many formats with a variety of deletion methods, and starting points. Wonghiransombat (2013) asserted that because of the weaknesses of the C-Test, a variation of the C-Test, the X-Test, was presented in 1987, which has more strengths than weaknesses.

The principle of the X-Test structure is based on the psychological perspective of Goodman (1967; cited in Boonsathorn, 1988), which states that for reading comprehension, the readers concurrently involve in all levels of processing—graph phonic, syntactic, and semantic. Unlike the C-Test, the X-Test deletes the first half of every second word. The deleted parts cover the semantics, while the remaining parts deal with syntactic/structural information.

SLA literature asserted that there are some studies in second language reading comprehension that are related to finding out the cognitive strategies involved in C-Test taking (Wonghiransombat, 2013; Stemmer, 1991, 1992; Boonsathorn, 1988), however, enquiries on the mental processes engaged in an X-Test seems to be so restricted.

For instance, for nearly a decade, the Bochum project on C-Test, focused on finding out the cognitive processes going on a student when he/she engaged in a C-Test. In order to discover “how” the learners solve C-Tests, subjects investigated in Bochum project were collected by using introspective and retrospective approaches and a model of analysis that focused on the assumption that C-Test is a mental task. The results showed that the C-Test is a cognitively demanding task (Cited in Grotjahn & Stemmer, 2002). Moreover, Stemmer (1991, 1992) tried to obtain more direct access to cognitive processes in C-Test taking by using think-aloud protocol. Her study’s findings do not only show some cognitive strategies involved in C-Test taking but also reveal differences in the effectiveness of strategies regarding different texts. Besides, Rahimi & Saadat (2005) revealed the mental processes of test-takers of the C-Test.

Furthermore, Babaii & Ansary (2001) in their study based on think aloud analysis, asserted that C-test tap macro level features and micro level features while micro-level features appear more often. Rahimi & Saadat (2005) based on think aloud protocol found that “subjects tend to use the bottom-up strategies quite more frequently than top-down ones for restoring the item” all of which to a significant degree depends on the text content, the mutilated words and proficiency levels of participants. In addition, based on Salehi, & Bagheri Sanjareh (2013) findings C-test and Cloze test trigger both macro-level strategies and micro-level strategies even though micro-level strategies were more appeared compare to macro-level.

Boonsathorn (1987; cited in Boonsathorn, 1988) paralleled the C-Test with the X-Test to discover the strategies
which L2 learners used in answering the C-Test and The X-Test. The results indicated that the C-Test and The X-Test were different functionally and structurally, and it seemed that the X-Test was more difficult and discriminated better than the C-Test for both L1 and L2 subjects because more of the normal reading process is required for the X-Test than for the C-Test.

As it is said earlier, there are still gaps in the literature regarding learners’ mental processes involved in an X-Test taking. In view of that, the current study investigated the possible use of cognitive strategies used by respondents, when they are engaged in answering the X-Test, to develop a better understanding of their thinking while participating in problem solving and to discover the true nature of what an X-Test measure.

The results of this study showed that each individual participant used a set of strategies linked to his/her cognition. These strategies recommend that the way people think and practice cognitive strategies may differ to each other which accordingly assert that students may act differently in problem solving processes, and this is the reason that Grotjahn & Stemmer (2002) stated, we cannot realize what a language test assess without an understanding of each test taker’s cognitive processes and mental operation on which the observed scores depend.

Moreover, examination of the data cleared that respondents engaged in using some kind of cognitive strategies for filling out the gaps of the X-Test. Specially, the strategy associated with making prediction by the use of co-text which was highly emphasized during the study. This is in accordance with previous studies available in literature regarding C-Test and Cloze test (Wonghiransombat, 2013; Stemmer, 1991, 1992; Boonsathorn, 1988; Rahimi & Saadat, 2002; Babaii & Ansary, 2001; Salehi & BagheriSanjareh, 2013). However, the extent to which each strategy was used, was significantly different. Participants also stressed that X-Test was more challenging than the C-Tests that they previously had taken and they had to think more deeply, that is in line with Boonsathorn (1987, cited in https://www.ukessays.com) that stated the ESL learners taking the X-Test needed more strategies than while they taking the C-Test. It is also in agreement with the study of Prapphal (1994) that showed the X-Test looked to be more closely connected to the cognitive and academic skills than the C-Test.

In view of that, and based on the existing literature, it seems that an X-Test has a better discriminating power than its original version namely C-Test (Cleary, 1988). Accordingly, Mehrpour (2012) proved that the order of deletion of letters in a test, whether the right part of the words are deleted (C-Test) or the left part (X-Test) affects the comprehension of the Test. Thus, it seems reasonable that students may use different strategic solution in X-testing versus C-test taking. He concluded that the comprehension of a C-Test was comparatively much easier than the comprehension of an X-Test.

To conclude the findings regarding the study’s research question, which is very new in EFL and ESL literature, researchers should assert that participants use different cognitive strategies during the X-Test taking, namely “Making prediction”, “Summarizing”, “Using connection”, “Using fix-ups”, “Making questions”, “Identifying a problem”, and “Reflecting”. However the frequency of using each strategy was different to the others. The findings revealed that participants used “Making prediction” as the most frequently used strategy by 61 percent of usage, and “Identifying a problem” as the least frequently applied strategy by 2 percent of usage.

Moreover, it is also revealed that the X-test was more challenging for the students comparing to their previous C-Test taking experience, as one of the students acknowledged that “In c-test we can see the beginning of the words that helped us in finding the word”, that is in line with the principles of reduced redundancy as indicated by Spolsky (1968, 1969).

6. Implication and Suggestion for Future Studies

This study proposes that think-aloud provides a wealth of information about learners’ thinking and discriminate them during problem solving tasks. Since think aloud protocol has intensified our understanding of what might occur inside of the participants during X-Test taking, its integration could provide a more complete insight into problem-solving behaviors. Furthermore, understanding the cognitive processes underlying the test aids researchers and students realize what the test actually measures. It also helps us in using the X-Test for educational purposes as a task. The study also boosts the use of multiple approaches for assessing learners’ reading skills due to individual differences in the use of cognitive strategies. The study can contribute to teachers, learners, researchers and even syllabus designers’ understanding of why a test taker has answered an item or why he or she has failed to do so. Furthermore, the study encourages further work in this area to investigate the other factors associated with filling the gaps of an X-Test, especially regarding to participants’ gender and their level of proficiency.
References


Spolsky, B. (1968). *What does it mean to know a language, or how do you get someone to perform his competence?* Paper presented at the second conference on Problems in Foreign Language Testing, University of South California.


**Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).