Introduction

In the last three decades, an important shift has been taken place in the field of second/foreign language teaching and learning, and researchers have focused mainly on learners' individual factors. Such focus includes the studies of learners' individual language learning strategies. Because as Chamot et al. (1999) point out that “Differences between more effective learners and less effective learners were found in the number and range of strategies used”. In addition, there is a relationship between the frequent use of learning strategies and achievement in the language (Green & Oxford, 1995; Oxford & Burry-Stock, 1995; Bremner, 1999). In such a way, since the publication of seminal works “What good learners can teach us” (Rubin, 1975) and “What can we learn from good learners” (Stern, 1975), there have done much valuable works in the field of Language Learning Strategies (LLSs).

To our knowledge, a review of the relevant literature considering LLSs shows that after decades of research in the related field, LLSs have received considerable attention in the literature. The researchers came to conclusion that a vast number of strategies have been reported to be used by language learners (Cohen,1990) through various used methods such as survey tools and written questionnaire (Gu & Johnson, 1996; Fan, 2003), interview (Gu, 2003; Parks & Raymond, 2004), think-aloud or verbal reports (Goh,1998; Nassaji, 2003), diaries or dialogue journal (Carson & Longhini, 2002), and re-collective narratives (Oxford et al.,1996). Such measurements are used in the single form of method (separately) or as component methods (single set of methods) based on nature and goals of research works.

The SILL, as a survey tool to measure LLSs, plays an important role in the field of English as a second/ foreign language. Based on the earlier research, the researcher aims to check its psychometrics in order to find whether it can be as a standard measurement when it is translated and adapted into Persian language.

Abstract

The current study aims to analyze the psychometric qualities of the Persian adapted version of Strategy Inventory for Language Learning (SILL) developed by Rebecca L. Oxford (1990). Three instruments were used: Persian adapted version of SILL, a Background Questionnaire, and Test of English as a Foreign Language. Two hundred and thirteen Iranian female university level learners of English language as a university major were the volunteers to participate in this research work. The intact classes were chosen. The psychometric qualities were tested. All the analyses conducted demonstrated that the SILL was equally valid and reliable to be used in Persian language to measure the Language Learning Strategies.

Keywords: Strategy inventory for language learning, Language learning strategy, English learning

The importance of language learning strategies

Since 1975s, dozens of various studies have contributed to our understanding of strategies employed by SL (Second language)/FL (Foreign language) learners at the level of adults. Such studies show in order to effect changes in perceptions of the learners’ role in the learning process, we need to discover more about what learners do to learn successfully. Moreover, such studies show that the best way of going about teaching strategies remain a subject of much debate right up to the present (Brown, 2001) because of use of LLSs is influenced by number of factors (Oxford, 1990), and the frequency and variety of LLSs vary among different individuals and based on a number of variables (Chamot & Kupper, 1989).

Since the use of appropriate strategies allow learners to take more responsibilities for their own learning, LLSs are seen as particularly important in henchmen of autonomy (Dickinson, 1987). In such manner, one of the important objectives in the study of LLSs is to determine effectiveness of LLSs in learning second/foreign language (O’Malley & Chamot, 1990). Therefore, if learners use LLSs efficiently, they can learn by themselves and self-examine their own progress. Therefore, having such situation for LLSs can improve learners’ performance and enhance their abilities of language learning. In this way, Chamot (2005) comes to conclude that LLSs as procedure which facilitate learning task. However, the wilter of research, previous research has failed to reveal the significant importance of individual differences in second/foreign language learning (Griffiths, 1991; Ellis, 1985).

Classification systems of language learning strategies

Although a number of researchers attempted to develop a classification scheme of LLSs, one of the important general problematic issues is the issue of typologies that have been formulated for classification of LLSs. For instance, sometimes one strategy belongs to one category can be classified under another category.
(Johnson & Johnson, 1998), or as Cohen (1998) states that they “are not clear-cut”. In such way, if there has not been unanimous consensus on definition of strategies, the same can be said for their classification.

As Gould (1981) presents “Taxonomy is always a contentious issue because the world does not come to use in neat little packages”, and there is not a logical and well-accepted system for describing of strategy (Oxford, 1994). In this way, finding a particular classification of LLSs as a universal basic classification that can be as a LLSs’ complete classification system, what everybody agrees upon, is impossible. However, a considerable progress has occurred in this field regarding the classification of LLSs (Ellis, 1994), and even it is clearly to find comprehensive taxonomies (O’Malley & Chamot, 1990; Ellis, 1994; Oxford, 1990).

From point of view of extensive review of the literature, Oxford (1990) gathered literature on LLSs, and the Oxford's taxonomy is “perhaps the most comprehensive classification of learning strategies to date” (Ellis, 1994).

There are many significant differences between Oxford's taxonomy and other ones. For example, firstly, Oxford classifies heterogeneous strategies into more specific categories (Ehrman et al., 2003), secondly according to O'Malley & Chamot (1990), Oxford’s strategy classification is an inclusion of every strategy that has up to then been cited in the learning literature; and thirdly Oxford's taxonomy links individual strategies and groups of strategies with each of the four language skills (Oxford & Burry-Stock, 1995). In this way, Griffiths (2004) suggests Oxford’ classification system of LLSs can be as a useful base for understanding LLSs.

Based on the Oxford's taxonomy, the SILL (Version 7.0) was developed. The SILL is used with learners of English as a second or foreign language, and it consists of 50 items. It was tested among learners of English who are native speakers of different languages which include translated version of 23 languages such as Chinese, French, German, Italian, Japanese, Korean, Spanish, Thai, and Turkish (Oxford, 1990), and used in more than 120 dissertations and theses (Lan, 2005).

The Oxford's taxonomy includes “direct” and “indirect” strategies, which is as a fundamental feature of Oxford's taxonomy (Ghasedy, 1998). Direct strategies are classified into three sub-categories that are Memory Strategies (9 items), Cognitive Strategies (14 items), and Compensation Strategies (6 items). Indirect strategies include three sub-categories that are Meta-cognitive Strategies (9 items), Affective Strategies (6 items), and Social Strategies (6 items). Oxford (1996b) presents “each sub-scale would have an adequate number of items to facilitate more in-depth understanding of the learning strategies” (p. 3).

The comparison and analysis between Oxford's classification (1990) and other LLSs classifications show that Oxford (1990) developed the strategy classification of O'Malley and Chamot (1990), and expanded it to encompass 62 kinds of strategies. Moreover, she broke down the social/affective category of O'Malley and Chamot (1990) in two categories, Social and Affective Strategies. Currently, Oxford and her colleagues are developing a task-based questionnaire to complement the SILL (Oxford et al., 2004).

**Methodology**

Participants

The descriptive statistics are such type of numerical representation of participants (Brown, 1996). The sample drawn from the population must be representativ so as to allow the researchers to make inferences or generalization from sample statistics to population (Maleske, 1995). As Riazi (1999) presents “A question that often plagues the novice researcher is just how large his sample should be in order to conduct an adequate survey or study. There is, of course, no clear-cut answer”. If sample size is too small; it is difficult to have reliable answer to the research questions. If sample is too large its difficulty how to do the research. To leave a margin of about 20% for ineffectual questionnaires slightly bigger numbers were chosen. In this way, initially a total of two hundred and fifty Iranian female learners of English language as a university major at the Islamic Azad University Branches of three cities named Abadan, Dezful and Masjed-Solyman in Khuzestan province in the southwest of Iran, were asked to participate in this research work. It must bear in mind that number of participants may affect the appropriateness of particular tool (Cohen & Scott, 1996).The intact classes were chosen.

The chosen participants for this study were female students studying in the third year of English major of B. A. degree, ranging age from 19 to 28 (Mean=23.4, SD=2). Their mother tongue was Persian which is the official language of Iran, according to Act 15 of the Iranian Constitution.

The socio-economic status of participants, such as the participants’ social background and parents’ levels of education was controlled as well by a background questionnaire, based on some indicators such as the parents’ socio-educational background and occupation. The participants were matched as closely as possible for socio-economic background to minimize the effect of social class. Accordingly, the participants were classified as a middle class. Moreover, most of the participants from the Islamic Azad University in Khuzestan province, Iran, have middle-class and similar socio-economic background.

Because of the nature of this work (regarding the use of the ELLSs), a general English proficiency test was used to determine the language proficiency level of participants in English in order to minimize the effect of English language proficiency. As Jafarpour (2001) defines “the percent classification of subjects by the experimental test that corresponds to those by the
criterion” (Golkar & Yamini, 2007), top of subjects are 27% and bottom of subjects are 27% (Golkar & Yamini, 2007), the participant who were classified as intermediate subjects, were asked to participate in the current study. Instrumentation in the current study

Three instruments were used to gather data in the current study. They were as follows:

Strategy Inventory for Language Learning (SILL): The Strategy Inventory for Language Learning (SILL) is a kind of self-report questionnaire that has been used extensively by researchers in many countries, and its reliability has been checked in multiple ways, and has been reported as high validity, reliability, and utility (Oxford, 1996a).

Initially the SILL used in North America with foreign language learners by different researchers such as Ehrman and Oxford (1989) and Oxford and Nyikos (1989), and it has increasingly been used in the Asia-Pacific region by researchers such as Yang (1999), Hiso & Oxford (2002) and Griffiths (2003).

The SILL is a structured survey (Oxford, 1990), which according to Oxford & Nyikos (1989) the strategies which were included in the SILL were gathered from extensive literature review. In addition, Oxford (1996a) claims in general, SILL reliability has been high, and the reliability remains “very acceptable” (Oxford & Bury-Stock, 1995,). Moreover, Green & Oxford (1995) claim that its reliability by using Cronbach's alpha ranges from 0.93 to 0.95 depending on whether the survey is taken in the learner's own language or in the target language. In addition, its reliability was reported in many studies as highly reliable in the translated version of different languages (Oxford & Nyikos, 1989; Grainger, 1997; Park, 1997; Griffiths, 2002; Abu Shamis, 2004; Sharp, 2008; Szu-Hsin et al., 2006; Yang, 2007).

Regarding the validity of the SILL, Oxford and Burry-Stock (1995), claims that the all types of validity are very high. In addition, factor analysis of the SILL is confirmed by many studies (Oxford & Burry-Stock, 1995; Oxford, 1996a; Hsiao & Oxford, 2002). Moreover, several empirical studies have been found moderate intercorrelation between the items of six categories in the SILL (Oxford & Ehrman, 1995). In this way, as Ellis (1994) believes the Oxford’s taxonomy is possibly the most comprehensive currently available.

Test of English as a Foreign Language (TOEFL): TOEFL (Structure and Written Expression, and Reading Comprehension parts) as a general English proficiency test was used.

A Background Questionnaire: The socio-economic status of participants was controlled as well through a background questionnaire.

The procedure of adaptation of SILL

In adaptation of Instruments from one language to another in research works, some problems occur, such as the problem of translation one questionnaire to another language (Perera & Eysenck, 1984). The researcher, in the case of the SILL, translated it into Persian. However, to check the accuracy of translation, the back translation was used by three English teachers and three psychologists who are fully proficient in both languages (English & Persian). The items were corrected until full agreement among the translators was achieved, and the pilot study confirmed such translated items of the SILL. Moreover, the balance between spoken and written Persian was checked.

One must bear in mind that the translation/back translation is one of more effective ways to solve the equivalent concepts of translated and original version of one questionnaire (Behling & Law, 2000), and one researcher can be an ideal translator if she or he is fluent in target language.

The original inventory includes 50 items, but the adapted version includes 49 items which adapted for the current study. In adapted version of the SILL, one item was taken out. The item was deleted based on the feedback from participants in the pilot study. The omission was in the part of Cognitive Strategies including the deletion of item number 22 “I try not to translate word”.

Sample of the pilot study

The sample for the pilot study, as “A small-scale replica and a rehearsal of the main study” (Riazi, 1999), was selected as it represented the entire sample for participants who were asked to participate in the main study. Since the sample size in the pilot study ranges from 20 to bigger of 65 (Hinkin,1998), thirty and nine female students of English language as a university major at Islamic Azad University Branches of three cities named Abadan, Dezful and Masjed-Soleyman were asked to participate in the pilot study.

Data collection procedures in the main study

The data for the study described in this study was collected between September 2010 and November 2010 in Iran, at the Islamic Azad University Branches of three cities that are named Abadan, Dezful, and Masjed-Soleyman. The period of time to administrate each of stages was assigned based on the pilot study.

Stage One: At this stage, the participants were asked to answer TOEFL test. Approximately 80 minutes were taken to answer the test (The first week).

Stage Two: At the second stage, the respondents were asked to fill the adapted SILL. The respondents were asked to respond to the questions within 10-15 minutes. Alongside the adapted SILL, Background Questionnaire was administrated as well (The second week).

Data analysis

After data collection, the data was entered onto database (SPSS) to enable data analysis to be carried out.

Results, discussion, and conclusion

It is noteworthy to mention that the classification of strength of correlation is not well accepted among different researchers, and there are different
classifications such as the classification suggested by Cohen (1988), Delavar (2010) and Ghiasvand (2008). In the current study, the classification that was suggested by Cohen (1988) was chosen as a criterion to interpret and discuss about the strength of the correlation (Table 1).

Table 1. The classification suggested by Cohen, J (1988)

<table>
<thead>
<tr>
<th>Level of strength</th>
<th>Amount of the strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>r = .10 to .29</td>
</tr>
<tr>
<td>Medium</td>
<td>r = .30 to .49</td>
</tr>
<tr>
<td>Strong</td>
<td>r = .50 to 1</td>
</tr>
</tbody>
</table>

The correlation holding between scores of the Persian adapted version of SILL are shown in Table 2. According to Table 2, all categories were significant positively correlated with each other at the p<0.01 level. Accordingly, for the category of Memory strategies and each of categories of cognitive strategies, meta-cognitive strategies, affective strategies and social strategies, the level of correlation was found at medium level. For the category of memory strategies and the category of compensation strategies, the level of correlation was found at low level.

Table 2 shows that for the category of compensation strategies and each of categories of cognitive strategies, meta-cognitive strategies, and social strategies, the level of correlation was found at medium level. For the category of meta-cognitive strategies and each of categories of cognitive strategies and affective strategies, the level of correlation was found at low level.

According to Table 2, for the category of meta-cognitive strategies and each of categories of memory strategies, compensation strategies and affective strategies, the level of correlation was found at medium level. For the category of meta-cognitive strategies and each of categories of cognitive strategies, and social strategies, the level of correlation was found at strong level.

Moreover, Table 2 illustrates that for the category of the affective strategies and each of categories of memory strategies, cognitive strategies, meta-cognitive strategies, and social strategies, the level of correlation was found at medium level. For the category of affective strategies and compensation strategies, the level of correlation was found at low level. Based on Table 2, for the category of social strategies and each of categories of memory strategies, compensation strategies, and affective strategies, the level of correlation was found at medium level. For the category of social strategies and each of categories of cognitive strategies and meta-cognitive strategies, the level of correlation was found at strong level.

Table 3. Internal consistency reliability coefficient for the whole and six sub-categories of the Persian adapted version of SILL

<table>
<thead>
<tr>
<th>Sub-category</th>
<th>Alpha</th>
<th>Number of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory strategies</td>
<td>0.67</td>
<td>9</td>
</tr>
<tr>
<td>Cognitive strategies</td>
<td>0.79</td>
<td>13</td>
</tr>
<tr>
<td>Compensation strategies</td>
<td>0.58</td>
<td>6</td>
</tr>
<tr>
<td>Metacognitive strategies</td>
<td>0.81</td>
<td>9</td>
</tr>
<tr>
<td>Affective strategies</td>
<td>0.57</td>
<td>6</td>
</tr>
<tr>
<td>Social strategies</td>
<td>0.72</td>
<td>6</td>
</tr>
<tr>
<td>SILL (Whole scale)</td>
<td>0.89</td>
<td>49</td>
</tr>
</tbody>
</table>

Internal consistency reliability coefficient for the whole and six sub-categories of the Persian adapted version of SILL are shown in Table 3. The internal consistency reliability coefficient was found 0.89 for the SILL. Moreover, for TOFEL, it was found 0.80. The reliability coefficient indicated the degree to which the results on a scale can be considered internally consistent, or reliable for the whole scale, some of sub-categories, and TOFEL (Nunnally & Bernstein, 1994; De Vellis, 2003; Moemeni, 2007; Ghiasvand, 2008). Such findings of reliabilities for the four instruments confirmed the finding of reliabilities in the pilot study. The important point which must be considered it is that since the number of items affects the reliability, for some of sub-categories, the internal consistency reliability coefficient was found low.
The construct validity of the Persian adapted version of SILL was tested through confirmatory factor analysis. According to Table 4, in the beginning results of factor analysis, KMO was found 0.781. Such finding confirms the factor analysis. Bartlett’s test of sphericity was found 3238.263 which was significant (sig=0.000). In such way, Table 4 shows that the data has accepted to assign the factor analysis.

Table 4. KMO and Bartlett’s test

| Kaiser-Meyer-Okin Measure of Sampling Adequacy | 0.781 |
| Bartlett’s Test of Sphericity | Approx. Chi-Square | 3238.263 |
| df | 1176 |
| Sig. | 0.000 |

Moreover, 9 items are available in the category of memory strategies and factor loads were in the \( \lambda =0.23-0.511 \) range; 13 items are available in the category of cognitive strategies and factor loads were in the \( \lambda =0.36-0.56 \) range; 6 items are available in the category of compensation strategies and factor loads were in the \( \lambda =0.39-0.54 \) range; 9 items are available in the category of meta-cognitive strategies and factor loads were in the \( \lambda =0.22-0.61 \) range; 6 items are available in the category of affective strategies and factor loads were in the \( \lambda =0.39-0.47 \) range; and 6 items are available in the category of social strategies and factor loads were in the \( \lambda =0.37-0.60 \) range. One the examining the t values in the simpls output, the regression coefficient of all the items in the model were found to be significant.

Generally speaking, based on the findings concerning the reliability and validity analyses of the Persian version of the SILL, it may be said that the scale can be a valid and reliable way to measure learners’ LLSs which are used.

Limitations of the present study

All educational quasi-research deals with living human beings occur out of laboratory conditions have limitations (Gall et al., 2003). Regarding the current study, since the SILL is a self-report and single source of information, it is not clear whether the participants actively used the LLSs that they claimed to use. Their response may be just their beliefs and thoughts that they have about their use of strategies. In such way, in order to investigate students’ actual use of LLSs, it should be some research method to corroborate results of the SILL. For instance, the researcher must observe classes, use think-aloud procedure (introspection), interview, and so forth. Secondly, there may also have been some unclear points in the SILL itself such as “Never” to “Always” may have been fuzzy because the interpretation of these scales can change according to the context (Hatch & Brown, 1995), and the vagueness of wording has been another persistent problem in using questionnaire (Gu et al., 1995). Thirdly, the SILL may not be able to cover all the dimension of learners’ LLSs and there is not deep insight. The fourth issue, although the learners’ use of strategy is dynamic across times (Schmitt, 1997), the SILL made strategy use to be as a static variable.

Moreover, some questionnaires such the SILL, do not describe in detail how a student uses strategies in response to any specific language task.

Regarding sample of the participants, the sample may not be representative the population in general. However, generally, researchers often select a convenience sample in order to complete a study (Gall et al., 2003).

Furthermore, concerning second/foreign language proficiency, determining proficiency in language learning for speakers of other languages is not an easy endeavor, and has extensively been discussed by experts (Farhady, 1982; Bachman, 1990). For instance, some studies used proficiency test scores (Green & Oxford, 1995; Dreyer & Oxford, 1996) and the others used self-rating (Wharton, 2000).

Regarding statistical methods, there is an important issue in the statistical procedures, that it is the reliability estimates the internal consistency may not be appropriate to measure something that could fluctuate in short period of time which is common in the most of studies. The test-retest reliability measure is better indicator of reliability in such type of research. The second limitation is response biases. As it is known, there are three prominent types of response biases, which are social desirable response, acquiescence, and extremely response bias (Herk et al., 2004).

References

64. Rubin J (1975) What the ‘good language learners’ can teach us. TESOL Quarterly. 9(1), 41-51.