

An Investigation of Reliability and Validity of Strategy Inventory for Language Learning among Iranian University Students

Hamid Mohammadi¹ & Khodaverdi Alizadeh²

¹ MA in TEFL, English Department, Faculty of Humanities, Islamic Azad University, Ahar Branch, Iran

² PhD Candidate, English Department, Faculty of Humanities, Islamic Azad University, Ahar Branch, Iran

Correspondence: Hamid Mohammadi, English Department, Islamic Azad University, Ahar Branch, Iran. Tel: 914-660-8244. E-mail: hamid.mohammadi27@yahoo.com

Received: May 24, 2014

Accepted: June 3, 2014

Online Published: June 9, 2014

doi:10.5430/ijelt.v1n2p53

URL: <http://dx.doi.org/10.5430/ijelt.v1n2p53>

Abstract

This investigation examined the reliability and validity of Rebecca L. Oxford's (1990) Strategy Inventory for Language Learning (SILL) among Iranian university students. Sample pools of 320 (male and female) non-major English students were recruited for participation in the reliability phase. The main instrument of the study was questionnaire. Findings from the present study revealed that the SILL score were test-retest reliable, displaying excellent reliability (Pearson's correlation > 0.8), with total scores not being significantly different across administrations. Exploratory and confirmatory factor analyses were conducted on the SILL to investigate construct validity. Confirmatory factor analyses was run on the data where a six factor structure was found to be the best fitting model for the tool, thereby confirming that SILL measures a multidimensional construct. The authors conclude that the SILL was reliable as well as valid to assess the language learning strategies in Iranian university context.

Keywords: Strategy inventory language learning, validity, reliability, confirmatory factor analysis

1. Introduction

It is clear that some people learn language faster and more effectively, than others. Given that their facilities, budget, and so on are the same, what are the causes of such differences? It seems that these people have a knack for their language learning and have built their skill as well as confidence over time than who are rather stopped or slowed down. According to research, in second language learners use more metacognitive strategies than FL learners, since they have more opportunities to use the target language (Riley and Harsch 1999).

An overview of language learning/acquisition reveals that in the middle of 70s, arguments took radical direction and more learner-centered approaches and tried to modify the character of the learner who achieve a desired result, and who has been specified as good language learner, for help weak learners (Rubin, 1975; Stern, 1975; Naiman et al., 1978). Rubin (1957) defines LLSs as the techniques which a learner may use to learners do to facilitate the obtaining, storage, retrieval of knowledge, to regulate learning. On the basis her definition, she introduced one of the earliest classification of language learning strategies. It consisted of two major categories: strategies that may contribute directly to learning and those which contribute indirectly to learning. These two broad categories were further divided into three primary categories: learning, communicative and social strategies. During the 80s and 90s, researchers under the prevalence of cognitive psychology, (Cohen 1998, O'Malley & Chamot 1990) criticized the classifications of Rubin and observed that theories of second language acquisition lacked a precise description of strategic processing in SLA. Their attempt resulted in formulating LLSs in an information-processing theoretical model (Kudo, 1999).

The definitions of learning strategies have changed over time, from the psychological view learner strategy is defined as set of operation, action, step, or routine that learner use to facilitate learning. While, from socioculture view learner strategy can be defined as higher mental process, such thinking, reasoning and planning (Cohen, 1998). To arrive at a definition for learning strategies Griffiths (2013, P. 5) states that from an extensive review of the literature spanning more than 30 years, a definition is suggested of language learning strategies as activities consciously chosen by learners for the purpose of regulating their own language learning.

Since different properties for definitions of learning strategies exist among researcher, there is no general agreement over the classification of learning strategies (Oxford, 1990). But among them Oxford (1990), yielded the most comprehensive and detailed classification of learning strategies to date, known as the Strategy Inventory for Language Learning (SILL). She defines LLSs as specific actions or behaviors taken by the learner to make language learning more successful, easier, self-directed, effective, and transferable to new situations (1998).

Oxford (1992, p. 20) noted that her system "is based on the theory that the learner is a 'whole person' who uses intellectual, social, emotional, and physical resources and is therefore not merely a cognitive/ metacognitive information-processing machine". Oxford's classification scheme consisted of two classes, six groups, nineteen sets and a total of sixty-two strategies. Direct and indirect strategies are major branches of her classification. The three direct strategies were identified as memory, cognitive and compensation strategies. The three indirect strategies were also designated as: metacognitive, social and affective strategies.

The strategies which require mental processing of the language and deal directly with the language itself are called direct strategies. On the other hand, strategies which deal with the general management of language learning, where the target language was not directly involved are called indirect strategies. However, Oxford pointed out that in most cases, direct and indirect strategies support each other.

It is claimed that this classification provide information not only about the strategy type, but also about task type and context in which the strategy can be use. Thereby, it is systematic and strategy group are linked with various skill such as writing, reading, speaking and listening (Kayaoglu, 1997).

1.1 Reliability and Validity

The reliability of a tool indicates its ability to yield consistent results across varying settings. In other words, it is tied up with the repeatability or reproducibility of any assessment (Coaley, 2010). The investigation of reliability according to Bachman (2005) is challenging with the question, "How much of an individual's test performance is due to measurement error or to the factors other than the language ability we want to measure?" (P. 160). One of the most important kind of reliability is test-retest procedure. Its main advantage is that it permits the instrument to be compared with itself, thus avoiding the sort of problems that could arise with the use of another instrument (Kumar, P. 180).

The validity of a screening tool indicates the ability of the tool. In the classical view of validity, construct validity is one of three main types of validity, as well as content and criterion validity. Modern view defines construct validity as the over-arching concern of validity research, encompasses all other types of validity evidence (Cronbach & Meehl, 1995). Testers believe that it is a superordinate kind of validity to which internal and external validity contribute (Ebil & Frisbie, 1991). Bachman (1990) argued that two kinds of evidences that confirm construct validity are correlation evidence and experimental. Correlation evidence is kind of statistical procedures that try to establish the relationships among variable. Correlation evidence to construct validation usually uses both exploratory and confirmatory modes. In the first one, we try to examine the correlation among a group of measures in order to modify the abilities that influence performance on tests. In the latter, we try to either confirm or reject these hypohtheses. Factor analytic methods can help scientists to define their variable more precisely and decide what variable they should study and relate to each other in the attempt to develop their science to a higher level (Andrew & Lee, 1992). So, in assessing validity of SILL this way will be take by the researcher.

1.2 Significance of Study

During last decade, we have witnessed an explosion of research in language learning strategies in Iranian context. Assessing tool in this field is more important, since "The research tool provides the input to a study and therefore the quality and validity of the output, the findings, are solely dependent upon it" (Kumar, P.156). It is evident that SILL has been adjusted to the U. S sample. Consequently, Although it is essential to perform CFA of the SILL in various context (Hsiao & Oxford, 2002), this study is initial investigation to confirm underlying constructs of the SILL among university students in Iran.

1.3 Review of Related Work

Green and Oxford (1995) announced that, reliability of SILL for the ESL/EFL ranges from .86 to .91 when learners respond to the questionnaire in their second language (English). Reliability coefficient increase when learners respond in their L1 to .91 to .99. Whether administered in the subjects L1 or L2 the SILL has high reliability. Although, Robson and Midorikawa (2001) in their study with Japanese university found unacceptable rate for test-retest reliability with a common variance of 19.5 and 25.5 at the subcategory and main category level respectively.

Results of factor analyses for the SILL have been conflicted (Robson and Midorikawa, 2001; Hsiao & Oxford, 2002; Woodrow, 2005; Eldib, 2004; Green & Oxford, 1995; Yang, 1999). For instance, El-dib (2004) found eight rather than the nine that had been previously reported by a prior study conducted by El-Dib. The identified factors explained 42.10 of the variability among 50 items on the SILL. Factor analysis among university students in Japan resulted in 15 factors structure. In addition, a disorganized pattern of strategies item resulted when they tried to fit administrations in to a six category solution. In corresponding study, among university students in Puerto Rico, it resulted in nine factor structure by Oxford study.

1.4 Research Questions

Current study tries to answer the following two research questions:

Research question 1: Does result of test-retest reliability from this sample yield similar pattern?

Research question 2: Does the construct validity of the SILL match with the hypothetical results?

2. Method

2.1 Participants

Sample pools of 320 (male and female) in non-major English were recruited in the reliability phase, although 305 students completed the study. Since only data from participants whom had answered all questions on both occasions were used in the analysis, data from 15 participants were further excluded for the reason either missing responses or not taking part in two occasions.

2.2 Instrument

Oxford's (1990) Strategy Inventory Language Learning, translated into the Persian language to elicit information. It consisted of 50 items figured out in six group. Nine questions belonged to memory strategy, fourteen questions belonged to the Cognitive strategy, nine questions belonged to metacognitive strategy, Five questions belonged to affective strategy, and seven questions belonged to social strategy. This questionnaire adopted a five-point Likert scale rating system: "1" stood for "never true of me"; "2" usually not true of me"; "3" stood for "Somewhat true of me"; "4" stood for "Usually true of me" and "5" stood for "Always true of me". Students were asked to choose only one of them according to what they really did rather than what they think they should have done.

2.3 Procedure

In general terms, the procedure was carried out in two phases (pilot and main study) and university students in the center of Tabriz were participated in this study.

2.3.1 Pilot Study

To ascertain that tool can be administered without any difficulty in the target population, the researcher completed a pilot study with thirty students for considering usability and practicability of the SILL and to avoid any difficulty which will encounter during the main study.

2.3.2 Implications for the Main Study

It appeared that three weeks apart between two occasions were appropriate to avoid of memorizing the answer by students. All of students were eager to fill out questionnaire and any of them felt impatient during administration.

Some students suggested that test instruction in SILL should be corporate with oral assistance in order to avoid misunderstanding of the test instructions. Considering layout of test a few students suggested that the layout of the test should be set attractively to draw attention of the tastes.

2.4 Main Study

The second phase of study was conducted in March, 2014. Participants met with the researchers on two separate occasions with three weeks apart.

2.4.1 First Visit

During first meeting, 320 copies of SILL were handed out among participants. The distribution and collection of the questionnaire was administered by university staffs. The researcher mentioned to the participants to fill out both sections of the questionnaires, one part to gain demographic information another related to the strategies that the participants may have recently used. Students received oral instruction in Farsi on how to complete the questionnaire. In order to prevent them from consulting and talking to their classmates, each one of the students who finished the

questionnaire went out of class.

2.4.2 Second Visit

The researcher met with the participant three weeks following the first visit to complete the retest. The researcher announced the participants two to three days prior to meeting as a reminder of the next visit by their teacher. After the participant completed the SILL, the daily video vocabulary were rewarded to participant and was thanked for their cooperation in this study.

2.5 Analyses of Data

Data collected were encoded into statistical package for the social science (SPSS) and three different levels of analyses were undertaken.

Descriptive statistic was run to look whether variables distributed in normal distributions or not. The Kurtosis, Skewness value $\leq \pm 2$ is considered well (Tabachnik and Fidell, 1996). The test-retest reliability was analyzed through the comparison of SILL score of the two separate occasions using paired samples T-test. In this case, a non-significant result indicates adequate agreement between the two occasions, implying reliability.

Exploratory and Confirmatory factor analysis were applied to assess the construct validity of SILL and the comparison of the researcher's theoretical expectations of the tool with the performance of the tool itself. For the exploratory factor analysis, the component principle extraction method was used, with factor loadings restricted to those greater than 0.3, and equamax rotation applied. For this project, a p value ≤ 0.05 was considered statistically significant.

3. Results

As can be seen in table 1, the results of descriptive statistics confirmed the exclusion of items 5, 18 that loaded kurtosis value of -3.15 and -3.10 respectively from analyses of data. A kurtosis value near zero confirms a shape close to normal. A negative value, confirms a distribution which is more peaked than normal, and a positive kurtosis confirms a shape flatter than normal.

Table 1. Descriptive statistics

Strategies	Mean	SD	Skewness	Kurtosis
Mem 1	2.62	1.42	.25	-1.62
Mem 2	3.35	1.32	-.14	-1.02
Mem 3	3.71	1.17	-.56	-.13
Mem 4	2.56	1.38	.26	-.96
Mem 5	3.55	1.09	-.59	-3.15*
Mem 6	2.57	1.50	-.51	-.89
Mem 7	4.10	.96	-1.25	1.84
Mem 8	3.39	1.19	.63	-.21
Mem 9	2.91	1.46	.10	-1.15
Cog 1	3.97	1.11	.72	.03
Cog 2	3.77	1.17	-.59	-.09
Cog 3	3.42	1.39	-.11	-.98
Cog 4	4.25	1.07	-1.27	-1.66
Cog 5	3.71	1.35	-.57	-.55
Cog 6	4.10	.64	-.41	-.24
Cog 7	3.89	1.07	-.07	-.08
Cog 8	4.19	1.12	-1.11	.83
Cog 9	3.38	1.45	-.17	-3.10*
Cog 10	3.22	1.21	-.06	1.16

Table 1. Descriptive statistics (continued)

Strategies	Mean	SD	Skewness	Kurtosis
Cog 11	3.41	1.45	-.01	1.41
Cog 12	3.98	1.22	.08	.22
Cog 13	3.52	1.36	-.17	1.06
Cog 14	4.15	.82	.08	1.82
Cog 15	3.01	1.34	-.02	-1.05
Com 1	3.71	1.06	-.53	-.09
Com 2	4.44	.95	-1.45	.59
Com 3	4.04	.94	-.55	.15
Com 4	2.62	1.38	.21	-.93
Com 5	2.64	1.27	.29	-.95
Com 6	4.20	1.23	-1.12	-.62
Met 1	2.63	1.36	.65	-.42
Met 2	4.44	.95	-1.45	.59
Met 3	4.04	.91	-.55	-.15
Met 4	2.74	1.33	.33	-.73
Met 5	3.15	1.47	-.16	-1.29
Met 6	3.60	1.24	-.66	-.04
Met 7	3.02	1.38	-.07	-1.07
Met 8	4.26	1.08	-1.30	1.08
Met 9	3.64	1.50	-.43	-1.18
Affec 1	4.08	1.11	-.79	.04
Affec 2	2.92	1.37	.05	-1.12
Affec 3	3.12	1.30	-.07	-.93
Affec 4	2.72	1.52	.15	-.96
Affec 5	3.45	1.44	-.16	-1.32
Soc 1	3.81	1.27	-.46	-.03
Soc 2	2.66	1.47	.16	-.98
Soc 3	2.67	1.60	.41	-.79
Soc 4	2.18	.09	.08	1.28
Soc 5	2.49	1.30	.53	-.11
Soc 6	3.01	1.56	.01	1.07
Soc 7	4.03	1.20	-.62	1.02

Note.* Indicates a kurtosis value $>+/-2$

A closer look at individual factor as shown in table 1, revealed interesting observation. Only, factor 5 and 18 had resulted in skewness and kurtosis value $>+/-2$, so they excluded from the analysis.

3.1 Research Question 1

As shown in Table 2, Paired samples T-tests and paired sample correlations were calculated on the total SILL score, as well as on individual attribute scores, indicating that attribute scores were not significantly different across time.

Table 2. Reliability scores of strategies

Strategy	Mean	SD	Pearson's Correlatin	T-statistic
Memory Strategies				
Time 1	27.71	4.48	.79*	1.43
Time 2	27.92	4.59		
Cognitive Strategies				
Time 1	38.91	5.28	.81*	1.73
Time 2	38.73	5.15		
Compensation Strtegies				
Time 1	16.78	5.87	.85*	1.78
Time 2	26.54	2.53		
Metacognitive Strtegies				
Time 1	29.53	5.86	.92*	-.94
Time 2	29.47	5.54		
Affective Strategies				
Time 1	19.40	4.85	.87*	.164
Time 2	19.75	4.94		
Social Strategies				
Time 1	18.54	2.51	.84*	1.32
Time 2	18.79	2.59		

Note.* Indicates a value that is significant at $p < 0.05$

3.2 Research Question 2

The second research question of the study, examined the construct validity of the SILL and asked whether results from this sample yield a pattern of factors similar to the factor analysis results of previous research studies. For the exploratory factor analyses, the Kaiser-Meyer-Olkin measure was 0.756, which indicates the sample size used in the exploratory factor analyses was "good". Barlett's test of sphericity ($\chi^2 = 11.87$, $p < 0.001$) was significant, indicating that it was appropriate to do a factor analysis on the data.

Table 3. Kaiser-Meyer-Olkin and Bartlett measures

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.756
Bartlett's Test of	Approx. Chi-Square	1158.74
	Df	1132
	Si	0.00

Factor analysis with equamax rotation was applied and resulted a total of 6 factors above the Eigenvalue of 1.00, which accounted for 69.32% of the total variance. These factors were also identified by the scree plot as shown in figure 1. These interpretations were based on Hatch and Lazaraton's recommendation that "A loading of .30 or above is considered to be a substantial link of a factor and test" (p. 494).

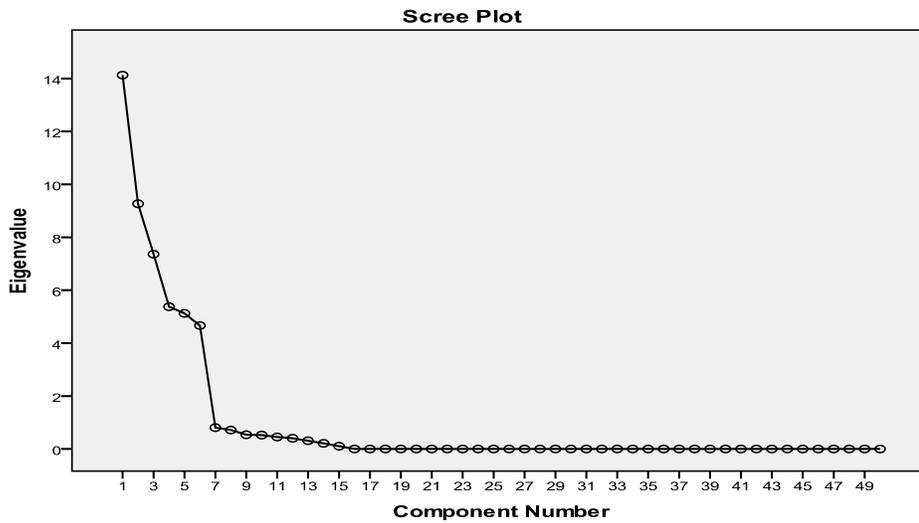


Figure 1. Scree plot for the SILL

Table 4. Principal Component Analysis, (n = 305)

Category	I	II	III	IV	V	VI
Mem 1	.49	-.04	.09	.12	.22	-.07
Mem 2	.81	.03	.02	.06	-.08	-.06
Mem 3	.57	.09	.08	.22	-.09	.12
Mem 4	.62	.19	.22	.11	.13	-.02
Mem 6	.65	.21	.03	.25	.08	.17
Mem 7	.49	.12	.10	-.16	.10	-.03
Mem 8	.41	.11	.01	.12	-.17	-.11
Mem 9	.43	.35	.04	.05	.23	.17
Cog 1	.34	.10	.20	-.06	.15	.13
Cog 2	.11	.39	.05	-.41	-.16	.15
Cog 3	-.10	.41	.09	-.07	-.01	.12
Cog 4	.13	.51	-.04	.14	.09	.05
Cog 5	.15	.78	.12	.21	.03	.06
Cog 6	-.14	.59	.25	.23	.04	-.10
Cog 7	-.22	.31	.17	-.12	.02	-.16
Cog 8	.06	.38	.08	-.09	-.09	.21
Cog 10	-.18	.73	-.01	-.16	.11	-.02
Cog 11	-.04	.53	-.13	-.08	.19	.21
Cog 12	-.28	.43	.02	-.16	.05	.02
Cog 13	.05	.32	.06	.07	.19	.11
Cog 14	.03	.52	.10	-.82	-.21	.15
Cog 15	.11	.44	-.08	.07	.19	.21
Com 1	.06	.38	-.11	.28	-.24	.19
Com 2	.07	.31	.13	-.19	-.09	-.18
Com 3	-.07	.05	.41	.04	.23	.02
Com 4	.11	.08	.32	.38	.12	.12

Table 4. Principal Component Analysis, (continued)

Category	I	II	III	IV	V	VI
Com 5	.06	.19	.59	.54	.16	.18
Com 6	.07	.01	-.06	.39	.15	-.05
Met 1	-.05	.28	.05	.34	-.22	.03
Met 2	-.07	.03	.52	.36	-.04	.02
Met 3	.11	.15	-.10	.54	.24	.03
Met 4	.14	.13	.11	.39	.05	.06
Met 5	-.17	.02	-.17	.49	.03	.09
Met 6	-.16	.21	.04	.51	.13	.15
Met 7	-.07	.14	.07	.09	.23	.03
Met 8	.11	.09	.42	-.15	.12	-.05
Met 9	.00	.03	.01	.51	.04	.13
Affec 1	.09	-.44	.04	-.09	.53	.21
Affec 2	.16	-.10	.01	-.05	.43	.12
Affec 3	.07	.21	.06	.02	.34	-.09
Affec 4	-.06	-.19	-.23	.21	.51	.01
Affec 5	-.02	-.08	.01	.04	.14	.31
Soc 1	.21	.21	.08	-.12	.09	.41
Soc 2	-.11	-.23	.14	.17	.18	.33
Soc 3	-.07	.03	.16	-.02	-.19	.36
Soc 4	-.18	.05	.02	.17	-.13	.43
Soc 5	.06	.19	.17	.23	-.09	.43
Soc 6	.13	.24	-.02	.03	.18	.51
Soc 7	.23	-.09	.19	.08	.12	.49

Note. I=Memory, II=Cognitive, III=Compensation, IV=Metacognitive, V=Affective, VI=Social

3.2.1 Confirmatory Factor Analyses

In the field of statistics, confirmatory factor analysis (CFA) is a kind of factor analysis. SILL was also tested using confirmatory factor analysis. It was conducted on, three, four, five and six proposed factors structures. Results of the confirmatory factor analyses can be seen in Table 5. Based on a comparison of the goodness-of-fit statistics produced from the confirmatory factor analysis, the six factors solution appears to be the most fitting for the SILL.

Table 5. Goodness-of-fit statistics

Goodness of Fit Statistic	3 Factor Structure	4 Factor Structure	5 Factor Structure	Six Factor Structure
χ^2	60.28	36.16	15.15	9.56*
GFI	.57	.62	.74	.82*
AGFI	.47	.78	.71	.82*
RMSEA	.30	.13	.13	.12*
CFI	.54	.84	.82	.86*

Note. * Indicates best fit indices

Acceptance or rejection of a model, depends on Indices of “goodness of fit”. Goodness of fit includes many type of fit measures that are crucial in interpreting the proposed model as following (see Table 5):

The first one is chi-squared index which commonly looks at discrepancies between observed and expected covariance matrices. It is Derived from the "fitting function" and usually affected by sample size. The best model confirms values near to zero. For the six factors model, it is estimated to be 9.56 for six structure model.

Goodness-of-fit statistic (GFI) is the second fit statistic reported. It was developed by Jöreskog and Sorbom and it is an absolute fit index that estimates the proportion of covariances in the sample data of the model. Values for this statistic range between 0.0 and 1.0 where 1.0 indicates the best fit. In the case of six factor structure, GFI is reported to be 0.82.

The next one, is adjusted goodness-of-fit statistic (AGFI) which usually increase with sample size. Values for the AGFI also range between 0 and 1 and it is generally accepted that values of 0.90 or greater indicate well fitting models. In the case of six factor structure, AGFI is reported to be 0.82.

Another commonly reported statistic is the root mean square error of approximation (RMSEA), as a badness of fit index introduced by Steiger and Lind (1980). An RMSEA in the range of 0.05 to 0.10 was confirmed a fair fit and values above 0.10 confirmed poor fit. For the six factors model, it is estimated to be 0.86 and suggested a best model fit.

The last investigated fit index is known as comparative fit index (CFI) compares performance on your model to performance on null model (Brown 2006). The CFI is one of a class of fit indices among the most commonly used in structural equation modeling. It is interpreted as the previous incremental indexes. For the six factors structure, it is estimated to be 0.12

4. Discussion

The current study, examined the test-retest reliability and construct validity of the SILL for use among Iranian university students. The test-retest involved recruiting of 320 Iranian university students. Data gathered from these samples were analyzed to investigate the test-retest reliability, and construct validity of SILL. In short, SILL found to be test-retest reliable, displaying excellent reliability (Pearson's correlation > 0.8), with total scores not being significantly different across administrations. Exploratory and confirmatory factor analyses were conducted on the SILL to investigate construct validity. The results from factor analysis, confirmed Oxford's six strategy categories. A number of 6 factors were resulted above the Eigenvalue of 1.00 which accounted for 69. of the total variance. The six factors were construed as follow: factor 1, as memory strategy that loaded 10 subcategories with the range of 34-81 value loading, Factor 2, as cognitive strategy that loaded 15 subcategories with the range of 31-78 value loading and also two subcategories in this group are loaded from compensation category unexpectedly. Factor 3, as compensation strategy that loaded 3 subcategories with the range of 32-59 value loading and also two subcategories from this group are cross loaded on the metacognitive category unexpectedly. Factor 4, as metacognitive strategy, that loaded 9 subcategories with the range of 34-54 loading. Factor 5, as affective strategy that loaded 4 subcategories with the range of 34-53 value loading and factor 6 as social strategy that loaded 8 subcategories with the range of 31-51 value loading, and 1 subcategories of this group are loaded from social category unexpectedly. Confirmatory factor analyses was run on the data where a six factor structure was found to be the best fitting model for the tool, thereby confirming that SILL measures a multidimensional construct. Lastly, the major finding of this research was that all the fit indices used to test the structure of the SILL such as Chi-square, GFI, AGFI, RMSEA and CFI were suitable to the data.

5. Limitation of Study

However, in conducting this research, certain limitations have appeared, and the fields for possible future research works should take these limitations into consideration: although the results of this study are useful to other university setting, they are non-inferred to other context because it is a case study of one university in Iran. Another limitation associated with measuring the test-retest reliability of the SILL. Since the participants completed the SILL on two occasions, three weeks apart, both the learning and memory effects may have affected the results of the study.

6. Recommendations

Current study is the initial investigation on Oxford's language learning strategies schema that performed in the context of Iran. Therefore, its results need to be confirmed. It would be valuable in replication of the present study among students in English major students. In addition, further research can be done to investigate other aspects different types of validity and reliability as well. This research focused on only construct validity and test-retest reliability of the SILL at Iranian university students

References

- Andrew L. Comrey, & Howard B. Lee. (1992). *A first course in factor analysis*. New York: Psychology Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Brown, D. (1978). *Principles of language learning and teaching*. Englewood Cliffs, NJ: Prentice Hall.
- Burry stock, J-A., & R. Oxford. (1993). *New analyses on expanded Puerto Rican strategy data*. (Unpublished manuscript), University of Puerto Rico at Mayaguez and University of Alabama, Tuscaloosa, AL. <http://www.sciencedirect.com/science/article/pii/0346251X9400047A>
- Comrey, L., & B. Lee. (1992). *A first course in factor analysis*. Newjersey: Lawrence Publication.
- Chamot, A. U., & P. El-Dinary. (1999). Children's learning strategies in immersion classrooms. *The Modern Language Journal*, 83(3), 319-341. <http://www.carla.umn.edu/immersion/acie/vol5/Nov2001>
- Coaley, K. (2010). *An Introduction to psychological Assessment and Psychometrics*. London: SAGE Publications Asia-Pacific Pte Ltd.
- Cohen, A. (1998). *Strategies in learning and using a second language*. London and New York: Longman.
- Cohen, A. Weaver, S., & T-Y. Li. (1998). The impact of strategies-based instruction on speaking a foreign language. In A.D. Cohen, *Strategies in learning and using a second language*. London: Longman. pp. 107-156.
- Cronbach, L. J., & Meehl, P. E. (1955). *Construct validity in psychological test*. *Psychological Bulletin*, 52, 281-302. <http://psychclassics.yorku.ca/Cronbach/construct.htm>
- Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In Lantolf, James P. (Ed.), *Sociocultural Theory in Second Language Learnin*. Oxford: Oxford University Press. pp. 27-50.
- Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed). Englewood Cliffs, NJ: Prentice Hall.
- El-Dib, M. (2004). Language learning strategies in Kuwait: Links to gender, language level, and culture in a hybrid context. *Foreign Language Annals*, 37, 85-95. <http://onlinelibrary.wiley.com/doi/10.1111/j.1944-9720.2004.tb02176.x>
- Green, J. M., & Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29(2), 261-297. <http://onlinelibrary.wiley.com/doi/10.2307/3587625/pdf>
- Griffiths, C. (2013). *The strategy factor in successful language learning*. UK: British Library publication.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. NewYork: Newbury House Publishers.
- Hsiao, T. Y., & Oxford, R.L. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *The Modern Language Journal*, 86, 368-383. <http://onlinelibrary.wiley.com/doi/10.1111/1540-4781.00155/pdf>
- Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.
- Kojic-Sabo, I., & P. M. Lightbown. (1999). Students' approaches to vocabulary learning and their relationship to success. *Modern Language Journal*, 83(2), 176-192. <http://onlinelibrary.wiley.com/doi/10.1111/0026-7902.00014>
- Kudo, Y. (1999). *L2 vocabulary learning strategies*. Honolulu: University of Hawai' I, Second Language Teaching & Curriculum Center. <http://scholarspace.manoa.hawaii.edu/handle/10125/8951>

- Kumar, R. (2005). *Research Methodology: A step-by-step Guide for Beginners*. London: Pearson Education. pp. 160-170.
- O'Malley, J. M., & A. U. Chamot. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Oxford, R. (1993). Gender differences in second/foreign language learning styles and strategies. In J. Sutherland (Ed.), *Exploring gender*. Englewood Cliffs, NJ: Prentice-Hall. pp. 140-147.
- Oxford, R., & J. Burry-Stock. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the Strategy Inventory for Language Learning. *System*, 23(2), 153-175. [http://dx.doi.org/10.1016/0346-251X\(94\)00047-A](http://dx.doi.org/10.1016/0346-251X(94)00047-A)
- Oxford, R. et. al. (2004). Effect of the presence and difficulty of task on strategy use: An exploratory study. *International Review of Applied Linguistics*, 42, 1-47. <http://www.deepdyve.com/lp/de-gruyter/effect-of-the-presence-and-difficulty-of-task-on-strategy-use-an-yse82nz0ji>
- Politzer, R., & M. McGroarty. (1985). An exploratory study of learning behaviors and their relationship to gains in linguistic and communicative competence. *TESOL Quarterly*, 19, 103-124. <http://onlinelibrary.wiley.com/doi/10.2307/3586774>
- Riley, L., & Harsch, K. (1999). *Enhancing the learning experience with strategy journals: Supporting the diverse learning styles of ESL/EFL students*. Paper presented at the HERDSA Annual International Conference, Melbourne, Australia. <http://www2.hawaii.edu/~kentonsrl/riley-harsch-99>
- Robson, G., & Midorikawa, H. (2001). How reliable and valid is the Japanese version of the Strategy Inventory for Language Learning (SILL)? *JALT Journal*, 23, 202-226. <http://jalt-publications.org/jj/articles/2652-how-reliable-and-valid-japanese-version-strategy-inventory-language-learning-sill>
- Rubin, J. (1975). What the 'good language learner' can teach us, *TESOL Quarterly*, 9(1), 41-51. <http://www.jstor.org/discover/10.2307/3586011?uid=2&uid=4&sid=21104242928773>
- Stern, H. (1992). *Issues and options in language teaching*. Oxford: Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. New York: Allyn and Bacon.
- Wenden, A. (1991). *Learner strategies for learner autonomy*. U.K.: Prentice Hall. Wharton, G. (2000).
- Woodrow, L. (2005). The challenge of measuring language learning strategies. *Foreign Language Annals*, 38(1), 90-100. <http://onlinelibrary.wiley.com/doi/10.1111/j.1944-9720.2005.tb02456.x/pdf>.
- Yang, N-D. (1992). The relationship between EFL learners' beliefs and learning strategy use. National Taiwan University. (Unpublished manuscript). <http://www.sciencedirect.com/science/article/pii/S0346251X99000482>