The latent structure of a Japanese version of the Strategy Inventory for Language Learning: A data-driven approach using exploratory factor analysis

ISEMONGER, Ian

Abstract

The Strategy Inventory for Language Learning (Oxford, 1990), or SILL, is claimed to measure strategies employed by language learners, and its presence in the literature is significant. The instrument has been translated into a variety of languages including Japanese, for which there is more than one translation. One of these translations was the subject of a confirmatory study by Isemonger (2016) using confirmatory factor analysis (CFA) as the method in an a priori test of the model hypothesized by the original author. The evidence for this model was negative. The analysis reported in this study takes an a posteriori approach to the data, using exploratory factor analysis (EFA) as the method. The analysis explores what the latent structure of scores generated by the instrument may be, given the lack of fit with the model hypothesized by the authors. Half of the items did not appear in the two EFA solutions extracted, and the remaining items partially aligned with Oxford's hypothesized scales, but also indicated that latents, thus far unspecified by the author, may be present in the data.

Keywords: Strategy Inventory for Language Learning, SILL, Learning Strategies,
Exploratory Factor Analysis, Unspecified Latents

Introduction

A significant area of inquiry connected to individual differences in second language acquisition (SLA) over the past four decades has been the branch of research we would refer to as language learning strategies. The study reported in this paper covers one instrument designed to measure language learning strategies, namely the Strategy Inventory for Language Learning (SILL; Oxford, 1990), and represents an exploratory approach to data collected for a Japanese version of the instrument. A previous study (Isemonger, 2016) has already provided negative evidence, under a confirmatory analytical framework, for the model hypothesized by the author (Oxford, 1990), using confirmatory factor analysis (CFA) as the method. The analysis reported in this paper adopts exploratory factor analysis (EFA) as the method, in line with the exploratory agenda of the study. A background on the roots of inquiry into language learning strategies and the emergence of the SILL is presented before outlining the methodological rationale for this exploratory study.

Isemonger (2016) offered a synopsis of the more remote roots of research into language learning strategies, which covered the notion of personal agency in learning and the relative suppression or elevation of the importance of this notion in major intellectual schools of thought present in the twentieth century. Fundamentally, schools of thought informed by behaviorism had suppressed the notion of personal agency in learning, as an exigency of the theoretical position rather than out

of ideological intent, because learning was construed as essentially passive; that is, learning was driven by the stimulus rather than the internal workings of an active and agentive mind. An agentive mind has capacities which necessarily exceed the input or stimulus, and behaviorism did not subscribe to this position. While behaviorism had ascended to a position of relative dominance in psychology and adjacent fields in the 1940s and 1950s, with foreign language learning and the associated audio-lingual method being an example of an adjacent field, alternative views of mind were curated through this period by two versions of constructivism, namely those of Piaget (2001a, 2001b) and Vygotsky (1978, 1986). Both of these versions of constructivism emphasized a notion of mind as something constructed in the process of the child's adaptation to the external world, with the end product being agentive and self-regulated rather than passive and stimulus-regulated. The critical difference between the two theoretical positions was that while Piaget did not treat the child as solipsistically engaged with the world, he also did not reserve an explicit and elaborated theoretical space, in accounting for the formation of mind, for the social other (i.e. the guardian, teacher and so forth), or even for language itself. Vygotsky, on the other hand, put the cognitive regulation of the child by the other, and through language, at the center of his theoretical framework, so that the agency of the other precedes metacognitive agency and self-regulation. In fact, and under Vygotsky's conception, not only does regulation by the other precede metacognitive agency and mature cognitive self-regulation, it is the very source of it.

Aside from these two important curators of a notion of agentive mind, it was the radical mentalism of Chomsky which not so much curated an alternative notion of mind, as unseated the behaviorist notion of mind altogether. All of these developments created an intellectual environment for the exploration of why some people are better at learning foreign or second languages than others, because agency, and the different expressions of it, is variable and therefore a predictor of learning outcome. Had the notion of a passive and stimulus-driven mind prevailed through behaviorism, these questions may not have been asked, and may not have precipitated lines of research on what the learner can do to accelerate learning; other than allow more reinforcement from the stimulus, of course. Whether one considers Vygotsky's or Piaget's version of constructivism, or Chomsky's mentalism, the same observation is pertinent. All of these theoretical positions view the mind as active rather passive, and above all as agentive. This allowed for explorative lines of inquiry concerned with the very variable outcomes in second an foreign language learning to flourish, and one of the early areas to flourish was work on the so called 'Good Language Learner' (Naiman, Frohlich, & Todesco, 1975; Rubin, 1975; Stern, 1975).

In the 1980's, there was an increasing interest in strategies in foreign language learning, and how these might affect learning outcome. In fact, there was a certain amount of crossover in the transition from work on the Good Language Learner to Language Learning Strategies. Early work by Bialystok (1978), who three years later was to publish an important paper on conscious learning strategies in language learning (Bialystok, 1981), cited and incorporated work from the above mentioned work by Rubin (1975) and Stern (1975). The overall trajectory was an acknowledgement that some people were better second or foreign language learners than others, and that conscious learning strategies had something to do with this. Of course, as a line of inquiry like this develops, one of the first concerns is to develop an ontology or taxonomy of possible learning strategies so that researchers have an explicit source of reference for more causal and explanatory lines of

inquiry. The development of such is important, because in order to measure learning strategies, something essential for any empirical line of inquiry, a conceptual order of constructs and sub-constructs, which can be measured as scales and subscales, needs to be established. The task of producing this conceptual order is essentially a descriptive one. Rubin (1981) offered an early taxonomy and a similar, but separate, line of research offering same (Chamot & Kupper, 1989; Chamot & O'Malley, 1987; O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985a, 1985b) followed at around the same time.

While the outlining of a taxonomy of constructs around a certain phenomenon of interest is a necessary precursor to its measurement, it is also true that if a particular taxonomy is not ultimately supported with measurement tools designed to make observations in terms of it, it is unlikely that the taxonomy will gain traction in the wider literature and associated empirical research trajectory. Once companion instrumentation is offered for a particular taxonomy, the empirical work propagated through use of the instrumentation tends to establish the associated taxonomy in the literature. The offering of companion instrumentation in the form of the SILL along with Oxford's (1990) taxonomy is arguably the reason why this taxonomy has become the dominant one in the field. The instrument and the taxonomy which underlies it has continued to occupy a significant space in the literature surrounding language learning strategies up to the present (e.g. Anam & Stracke, 2016; J. Green & Oxford, 1995; Hong-Nam & Leavell, 2006; McCullen, 2009; Nisbet, Tindall, & Arroyo, 2005; Sun, Mantero, & Summers, 2014; Tam, 2013; Watson & Ebner, 2018; Wharton, 2000; Yang, 1999).

The significant presence of an instrument in the literature should not, however, be read as evidence for the validity and reliability of scores it produces and, in fact, Wilkinson and the APA Task Force on Statistical Inference (1999), almost two decades ago, cautioned on the propensity for instruments to gain traction in the literature despite an absence of evidence for being able to produce structurally valid scores. With respect to the SILL, an early examination of the psychometrics of scores generated by the instrument would include a study by Robson and Midorikawa (2001). This study found respectable Cronbach's alphas, but also EFA solutions which were not aligned with the structure hypothesized by Oxford (1990). A study by El-Dib (2004) a short while later, also using EFA, produced a similar outcome with numbers of factors and items on factors not aligned with the structure hypothesized by the author. Much more recently, results have been obtained by Petrogiannis and Gavriilidou (2015) which indicate that the problems are persistent and difficult to overcome. These authors employed EFA to arrive at a subset of 29 items (thus 21 items were discarded), but with these still structured within the six constructs hypothesized by Oxford. A subsequent CFA was conducted by these authors using the same data as was used in the EFA, therefore meaning that the CFA they conducted was not an a priori test and was capitalizing on chance. However, despite this advantage for the test, values for critical indexes such as the TLI, CFI and NFI were all well below the threshold of >.95 (the reported values for the total sample and subsamples were generally around .80 and in no case higher than .86). These indexes are not test statistics, and therefore are interpreted on a continuum with some latitude for not meeting the threshold, but their reported values were, nonetheless, very far below the threshold.

Isemonger (2016), a year later, also conducted a CFA, using the model hypothesized by Oxford

(1990); and unlike the Petrogiannis and Gavriilidou (2015), the test was an a priori one. The CFA used a sample of 756 freshmen at a university in Western Japan, and the model tested was Oxfords six-construct model with all 50 items included. The evidence for the model was negative. While the Cronbach's alphas reported in this study were respectable (ranging between .70 and .88), it is the results from the CFA which were the most important. Cronbach's alpha is widely reported, and widely perceived as evidence for structural validity of scores, but it is in fact an index of reliability with the critical limitation that it does not demonstrate unidimensionality of scores on a scale or subscale (Cortina, 1993; S. B. Green, Lissitz, & Mulaik, 1977). The results for the TLI and CFI indexes (the two incremental fit indexes selected for the CFA) were .76 and .77, respectively. Both of these results are supposed to be greater than .95 with some latitude for interpretation around the threshold, because these are indexes and not test statistics. The result for the RMSEA (.06) was acceptable. It should be less than .60, but results on, or a little higher than, the threshold can be interpreted as acceptable given that this is an index. The RMSEA rewards model parsimony, i.e. reducing many items to few constructs, and the SILL is parsimonious with 50 items reducing onto only 6 constructs. The residuals were also low (SRMSR) at .06, meaning this result was acceptable. Taken in triangulation, however, which is how the results for these indexes are supposed to be taken (Hu & Bentler, 1999), the evidence for the model was negative.

When a direct and a priori test of the model hypothesized for an instrument by its author is rejected, the research question which naturally follows concerns what the latent structure of scores generated by the instrument actually is, if not in alignment with the structure advocated by the author. This kind of question is best pursued in an exploratory manner using EFA as the method, and is essentially an a posteriori approach to the data. EFA assists with finding out, for example, how many factors should actually underpin the scores (with this number led by the data and not by how many the author says there are), which items are unhelpful for these factors (i.e. do not indicate any of them), and which items are miss-specified (i.e. represent a factor other than the one they were expected to represent). EFA also assists with establishing whether there may be as yet unspecified latents lurking in the data, or put another way, whether there are hidden constructs represented by scores on certain items which are not included in the model claimed by the author. The purpose of this study was, therefore, to use the same dataset in which the a priori test failed (Isemonger, 2016), and use the alternative method of EFA to explore what the structural properties of the scores generated by this Japanese version of the SILL might be, if not in alignment with what Oxford (1990) hypothesizes.

Method

The analysis conducted in this study involved an exploratory rationale using EFA as the method. The overall method is reported below in terms of the instrument itself, the participants, and the procedure.

Instrument

The Japanese version of the instrument used to generate scores for this analysis (and the previous

analysis in Isemonger [2016]) is already present in the literature (see for example, Iwasaki, 2006). Information on the procedure for its translation is not available in the literature. However, and despite this deficit, given its presence in the literature, the latent structure of scores it produces is of research interest, because research is premised upon it. It must also be presumed that the instrument has entered professional practice, and therefore the structural validity of scores it produces is of interest to the professional and teaching community.

The original instrument authored by Oxford (1990) has 50 items, and these items are organized in a sequential rather than random manner on six interpretive subscales including: Memory Strategies (Items 1 - 9), Cognitive Strategies (Items 10 - 23), Compensation Strategies (Items 24 - 29), Metacognitive Strategies (Items 30 - 38), Affective Strategies (Items 39 - 44), Social Strategies (Items 45 - 50). The instrument is responded to on a Likert scale which has the typical five points indicating degree of endorsement of the item with respect to self. It is clear from findings by Isemonger in the 2016 paper, cited and discussed above, that the latent structure for scores, generated by the Japanese version used, did not fit the model specifying the above six constructs.

Participants

A total of 837 participants were involved in the survey. Initial inspection of the data revealed that 81 of these 837 cases contained missing data; usually, but not always, a single item, probably omitted by mistake. The missing data was observed to be at random rather than systematic and so these 81 cases were removed, leaving a final dataset of 756 cases. The participants were all starting their first year of study, and given the homogeneous age demographic of students accepted into Japanese universities, it was expected that almost all of the participants would be close to 19 years of age. This was the case, and as previously reported in Isemonger (2016), 99 percent of the sample fell between 18 years and 21 years. In terms of gender, the sample was relatively well balanced with 389 (52%) males and 365 (48%) females.

Procedure

Participants were invited to complete the SILL, but were informed that it was not required. All data was collected under informed consent. Administration of the SILL took between 15 and 20 minutes with few exceptions; but at no point were students under the impression that there was a time limit for completing it. The analytical procedure for the initial confirmatory study was reported in Isemonger (2016), and these results failed to confirm the model hypothesized by the original author. This finding invited the question as to what the latent structure for scores actually is, if not in empirical conformity with the model hypothesized by the author. This kind of question is essentially exploratory and the analytical procedure for this involves EFA. The data was, therefore, and for the purposes of this study, subjected to EFA for which two solutions were derived, with these and their associated rationales being reported below.

Results

The central contribution of the research reported in this paper is the EFA. The solutions pursued in the EFA analyses represent new analyses of the data, and from a bottom-up, or data-driven,

perspective. The descriptive statistics were briefly reported in Isemonger (2016), but are reported again below, under citation, and with more detail concerning the distributional properties of the scores for every item, because this higher-resolution information is interpretively useful when accounting for the solutions derived for the EFA.

Descriptive Statistics

The descriptive statistics are available for inspection in Table 1, and the result for the mean, standard deviation, skew (and the associated critical ratio [CR]) and kurtosis (and the associated critical ratio [CR]) is reported for each item.

Table 1
Means, Standard Deviations, Skew (and Critical Ration [CR]) and Kurtosis (and Critical Ratio [CR]) for Scores on all 50 Items Comprising the SILL

	Mean	Std. Dev	Skew	CR	Kurtosis	CR
Item 01	2.747	0.996	0.264	*2.963	−0.171	−0.959
Item 02	2.495	0.962	0.439	**4.928	−0.016	−0.090
Item 03	2.798	1.080	0.130	1.461	−0.632	**−3.549
Item 04	2.669	1.062	0.370	**4.149	−0.455	*−2.552
Item 05	2.578	1.104	0.322	**3.613	−0.602	**−3.376
Item 06	*3.239	1.216	−0.213	*−2.391	−0.869	**−4.875
Item 07	2.070	0.973	0.877	**9.846	0.438	*2.459
Item 08	2.495	0.903	0.357	**4.002	0.047	0.264
Item 09	2.713	1.113	0.185	*2.071	−0.721	**−4.046
Item 10	*3.765	1.049	−0.609	**−6.837	−0.158	−0.885
Item 11	2.295	1.042	0.669	**7.511	−0.087	−0.486
Item 12	2.857	1.104	0.225	*2.528	−0.629	**−3.531
Item 13	2.348	0.933	0.598	**6.718	0.151	0.847
Item 14	1.804	0.828	1.009	**11.323	1.027	**5.761
Item 15	2.534	1.145	0.395	**4.431	−0.600	**−3.369
Item 16	1.616	0.860	1.498	**16.817	2.096	**11.765
Item 17	1.488	0.754	1.712	**19.218	3.120	**17.514
Item 18	*3.030	1.178	0.004	0.048	−0.805	**−4.519
Item 19	2.472	1.088	0.469	**5.265	−0.336	−1.887
Item 20	2.739	1.071	0.170	1.904	−0.502	*−2.815
Item 21	2.966	0.920	0.068	0.764	0.096	0.541
Item 22	2.901	1.058	0.192	*2.151	−0.518	*−2.909
Item 23	2.553	0.959	0.372	**4.172	−0.061	−0.342
Item 24	*3.553	0.957	−0.115	−1.288	−0.352	−1.977
Item 25	*3.516	1.116	−0.320	**−3.592	−0.640	**−3.592
Item 26	2.675	1.111	0.348	**3.903	−0.532	*−2.984
Item 27	1.996	0.888	0.735	**8.253	0.334	1.873
Item 28	2.093	0.922	0.730	**8.191	0.437	*2.455

Item 29	*3.176	0.991	−0.104	−1.169	−0.135	−0.756
Item 30	2.533	0.948	0.507	**5.696	0.078	0.439
Item 31	2.745	1.000	0.297	**3.333	−0.168	−0.943
Item 32	*3.481	1.002	−0.245	*−2.752	−0.204	−1.147
Item 33	2.733	0.986	0.338	**3.799	−0.155	−0.872
Item 34	1.873	0.846	0.980	**11.005	1.252	**7.024
Item 35	1.993	1.086	1.111	**12.472	0.675	**3.790
Item 36	2.200	1.068	0.733	**8.226	−0.011	−0.061
Item 37	2.452	1.234	0.632	**7.100	−0.519	*−2.911
Item 38	*3.204	1.107	−0.074	−0.828	−0.554	**−3.110
Item 39	2.705	1.022	0.294	**3.296	−0.251	−1.410
Item 40	2.517	1.000	0.553	**6.211	0.143	0.801
Item 41	2.612	1.076	0.433	**4.861	−0.292	−1.637
Item 42	2.935	1.081	0.129	1.446	−0.544	**−3.051
Item 43	1.329	0.627	2.227	**24.998	6.074	**34.091
Item 44	2.118	0.998	0.826	**9.269	0.381	*2.136
Item 45	*3.352	1.070	−0.138	−1.547	−0.544	**−3.052
Item 46	2.286	0.988	0.635	**7.126	0.064	0.359
Item 47	1.917	0.882	0.753	**8.455	0.131	0.734
Item 48	1.976	0.929	0.870	**9.771	0.533	*2.991
Item 49	2.015	0.908	0.831	**9.328	0.662	**3.713
Item 50	2.690	1.129	0.296	**3.326	−0.482	*−2.703

Note

1. The asterisks on results for items in the mean column indicate means which fell above 3.0 (only nine items).
2. The single asterisks on results for items in the skew CR and kurtosis CR columns indicate values which fell above the strict threshold of 2.0 but below the more relaxed threshold of 3.0.
3. The double asterisks on results for items in the skew CR and kurtosis CR columns indicate values which fell above the more relaxed threshold of 3.0.

With respect to the means for the 50 items comprising the SILL, the most important observation is that these means predominantly fall below the mid-point of 3.0 on the scale. There were only nine items which presented with means above 3.0, and no items where the mean fell above 4.0. This suggests a general pattern of responses which veer away from endorsing the statement in an item as representative of self.

With respect to score distributions, and as previously reported by Isemonger (2016), there were only 15 items which met a (more relaxed) threshold of 3.0 for the critical ratio for skew. In this analysis, it was also decided to introduce a stricter threshold of 2.0, and in the case of this threshold only nine items were acceptable. An additional observation on skew, and one which proceeds beyond the previous study, is that there are numerous cases of extreme skew; in a few cases going beyond Critical Ratio = 10.0, and in one particularly serious case, even beyond Critical Ratio = 20.0. In the case of kurtosis, the overall picture was better with 22 items meeting the stricter threshold of 2.0. However, there were a few serious cases of kurtosis, with Item 43 having

a value over 20.0, and this was the same item which presented with the most serious case of skew (also over 20.0).

Exploratory Factor Analysis

The executional sequence for the EFA followed the typical sequence of 1) first testing the suitability of the matrix for EFA, 2) extracting an initial un-rotated Principal Components Analysis (PCA) to assist with determining the number of factors to extract, and 3) extracting solutions informed by the information derived from the initial solution.

With respect to the first step, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was requested, and the value returned was .945. This was quite far above the threshold value of .80 typically adopted in the literature, indicating the matrix was suitable for EFA. Additionally, Bartlett's Test of Sphericity was executed, and the result was significant ($p < .01$). This also supported the position that the matrix was suitable for EFA.

With respect to the second step, an un-rotated PCA was executed with the number of components to extract unstipulated. This type of initial solution extracts factors until the matrix is exhausted, and the eigenvalues (associated with variance accounted for by each successive factor extracted) are returned. The eignenvalues can then be inspected under the eigenvalue-greater-than-one rule to determine the number of factors to extract. Alternatively, a scree plot (which is essentially a plot of the eigenvalues) can be visually inspected to determine at what point, in the extraction of factors, the eigenvalues start to level off.

The eigenvalues for the un-rotated PCA executed on the SILL data matrix can be inspected in Table 2. The eigenvalue dropped below 1.0 on the 11th factor suggesting that 10 factors should be extracted. However, the eigenvalue-greater-than-one rule has a tendency to overstate the number of factors which should be extracted (Zwick & Velicer, 1982, 1986), and furthermore, inspection of the scree plot (Figure 1) indicated a different picture, with a leveling-off after the fourth factor. It is very clear from the scree plot that although the fifth through tenth factors all have eigenvalues which are greater than 1.0, they also add about the same amount in each case to variance accounted for, and an amount which could be considered trivial.

Table 2
Table of Eigenvalues (for Un-rotated PCA), and Variance and Cumulative Variance Accounted for by Each Factor (up to the 11th Factor)

Component	Initial Eigenvalues		
	Total	**% of Variance**	**Cumulative %**
1	13.816	27.632	27.632
2	2.653	5.306	32.938
3	2.070	4.141	37.079
4	1.547	3.095	40.174
5	1.492	2.984	43.158
6	1.346	2.691	45.849
7	1.265	2.529	48.378

8	1.167	2.334	50.713
9	1.104	2.208	52.921
10	1.056	2.112	55.033
11	.999	1.999	57.032

Note
The eigenvalues are only represented in this table up to the 11th Factor where the value drops below 1.0. Subsequent factors were extracted up to the 50th Factor (equal to the number of items, and at which point the data matrix is necessarily exhausted), but these are not represented in the table.

The decision was therefore to extract a four-factor model. However, an initial execution of the four-factor model, under an oblique rotation, and using ML estimation, failed to converge. A more sophisticated procedure, namely parallel analysis, the rationale for which was first outlined by Horn (1965), was therefore adopted. The procedure is not built into SPSS, but has to be executed using syntax, now widely adopted and written by O'Connor (2000). In the case of this execution, the syntax for raw data was selected (with 1000 permutations) and using the 95th percentile. Using data permutations assists with non-normal properties in the data, and such properties were a feature of this dataset.

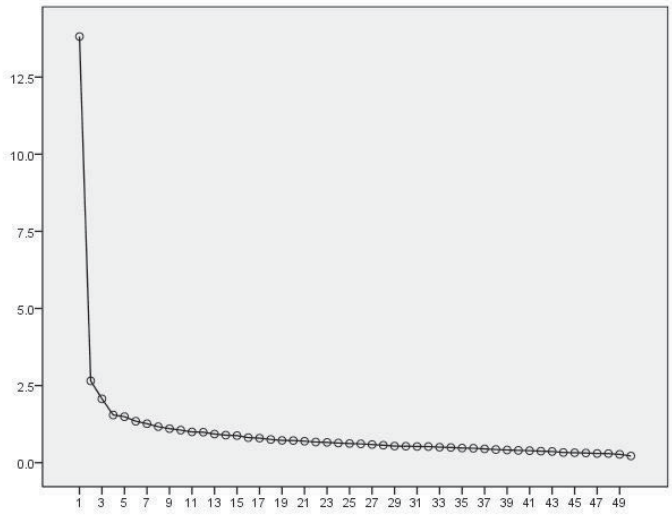

Figure 1. Screeplot of Eigenvalues for Each Factor Extracted in the Un-Rotated PCA. The y axis represents eigenvalues, and the x axis represents factors extracted by number.

According to the parallel analysis procedure, and in comparing the eigenvalues for the raw data to those for the 95th percentile for the random-data eigenvalues generated by the syntax, the number of factors to extract should be five. Therefore, a five-factor solution was extracted using ML estimation and under an oblique rotation (Direct Oblimin), and there were no problems with convergence (convergence occurred after 20 iterations). This five-factor solution is reported

immediately below.

In addition, a six-factor solution was also executed, because the original author (Oxford, 1990) hypothesizes this number of factors to underlie scores produced by the instrument. If five factors are extracted, it is inevitable that the original structure will not be maintained because the number of factors for items to reduce onto is less than hypothesized. Thus, although there was no empirical rationale (like the results from a parallel analysis) for extracting six factors, a six-factor solution acted as a point of reference to establish on which factors items would appear if the number of factors was left unchanged from the number hypothesized by the author.

Five-Factor Solution

The resulting coefficients for the factor pattern matrix for the five-factor solution are presented in Table 3. Twenty five items, or half of the total items, failed to emerge on any factor with a coefficient of above .40. Another important overall observation is that no items hypothesized to measure Affective Strategies appeared in the solution. Given that five factors were extracted, rather than the six hypothesized by Oxford (1990), the loss of one of the original constructs in the solution should not be seen as surprising, but this finding also points to the possible conclusion that the Affective Strategies scale may be one of the weakest in the instrument.

Table 3.
Factor Pattern Matrix for the Five-Factor Solution

		Factor				
		1	**2**	**3**	**4**	**5**
Item 01	Memory Strategies					
Item 02						
Item 03			.678			
Item 04			.524			
Item 05						
Item 06						
Item 07			.531			
Item 08						
Item 09						
Item 10	Cognitive Strategies					
Item 11						
Item 12						
Item 13						
Item 14				.503		
Item 15						
Item 16				.472		
Item 17				.551		
Item 18						
Item 19			.514			
Item 20			.536			

Item 21			.505			
Item 22					−.500	
Item 23						
Item 24	Compensation Strategies				−.581	
Item 25					−.506	
Item 26						
Item 27						
Item 28						
Item 29					−.618	
Item 30	Metacognitive Strategies					
Item 31						
Item 32						
Item 33		.630				
Item 34		.450				
Item 35		.688				
Item 36		.803				
Item 37		.820				
Item 38		.711				
Item 39	Affective Strategies					
Item 40						
Item 41						
Item 42						
Item 43						
Item 44						
Item 45	Social Strategies				−.482	
Item 46						.504
Item 47						.548
Item 48						.729
Item 49						.634
Item 50		.415				

Note

1. The second column indicates the construct which the items are hypothesized to measure by the author (Oxford, 1990). Values below .40 are suppressed in this table. Rows which are shaded indicate items which did not appear on any factor above a value of .40.
2. Labels assigned to each factor were as follows: Factor 1, Metacognitive Strategies (Revised 5F); Factor 2, Cognitive Strategies (Revised 5F); Factor 3, Authentic Language Use Strategies (Revised 5F); Factor 4, Tolerance for Communicative Imprecision (Revised 5F); Factor 5, Social/Communicative Strategies (Revised 5F).

Factor 1 included Items 33 through 38 with all of these items having been hypothesized to measure Metacognitive Strategies by the author (Oxford, 1990). Items 30 through 32, which were also hypothesized to measure Metacognitive Strategies, did not appear along with the above six

items, and did not appear on any other factors either. The label of *Metacognitive Strategies (Revised 5F)*[1] was given to this factor due to the relatively good showing of six of the originally-hypothesized items on the factor.

Factor 2 included three items from the author's original Memory Strategies scale (Items 03, 04 and 07) and three items from the original Cognitive Strategies scale (Items 19, 20 and 21). Ascribing a label to this factor was problematic. Items 03 and 04 directly mention remembering in their content, but could also be construed as cognitive in nature. Item 07 refers to physically acting out new words encountered in the second language, which seems to imply a kinetic aspect which Items 03 and 04 do not. Items 19, 20 and 21, from the original Cognitive Strategies scale, seem to superficially represent cognitive-type strategies. Disregarding Item 07, it was therefore decided to label this construct *Cognitive Strategies (Revised 5F)*, but with some empirical qualification and some interpretive caution.

Factor 3 comprised three items from the author's Cognitive Strategies scale (Items 14, 16 and 17) and no other items. These three items appeared to reference authentic language use as a strategy, and therefore Factor 3 was labelled *Authentic Language Use Strategies (Revised 5F)*.

Factor 4 comprised five items. These included one item from the author's Cognitive Strategies scale (Item 22) and three items from the author's Compensation Strategies scale (Items 24, 25 and 29). There was also a single item from the original Social Strategies construct (Item 45). All items on this factor had negative loadings (i.e. a negative orientation to the factor). In terms of ascribing a label to the scale, there were two possible rationales for a label. The first was to retain the label of Compensation Strategies (with the "Revised" suffix), because three of the items came from the originally-hypothesized Compensation Strategies scale, and these therefore dominated the factor. Items 22 and 45 which both appeared on Factor 4, but which did not come from the original Compensation Strategies scale, would then be treated as anomalous. The second rationale for a label seemed better, however, and this rationale focused on item content relating to a tolerance for communicative imprecision in the second or foreign language. This seemed better because such content was evident in the items, and it covered the content of four of the items quite well, with only Item 45 having to be treated as potentially anomalous. This factor was therefore labelled the *Tolerance for Communicative Imprecision (Revised 5F)* scale. The term "Strategies" was omitted for this label, because the item content is argued to reference a tolerance rather than a strategy.

Finally, Factor 5 included four items from the original Social Strategies scale (Items 46 through 49), and only two items from the original scale were not included. Item 50 (not included) referenced learning about culture. In this respect, it departed from the other five items which did not reference culture at all. Item 45 (also not included) appeared to be the only statement from the Social Strategies scale which used a conditional grammatical construction, rather than a simple affirmative sentence. These four items could therefore have retained the original label, but inspection of the item content indicated that they could just as easily fall under the label Communicative Strategies. Therefore, the final label ascribed reflects this ambiguity and was given as *Social/Communicative Strategies (Revised 5F)*.

---

[1] The term "Revised" in brackets indicates that while the name is the same as the original, the items comprising it are not. The "5F" indicates that the factor is revised within the five-factor solution.

Six-Factor Solution

As mentioned earlier, the six-factor solution has no empirical rationale, unlike the five-factor solution above, and is therefore not led by the data. It was extracted as a point of reference, because the original author (Oxford, 1990) hypothesized six constructs. Also, while there was an empirical rationale for extracting five factors via the results for the parallel analysis, it is also inevitable that one of the original constructs would fail to find representation on a five-factor solution; although, even more could fail to find expression depending on other contingencies in the data. The six-factor solution does not, as an inevitability, have this consequence; although it may be the case that some of the original factors are lost anyway, and again, depending on other contingencies in the data. The results for the six-factor solution are available for inspection in Table 4.

Table 4.
Factor Pattern Matrix for the Six-Factor Solution

		Factor					
		1	**2**	**3**	**4**	**5**	**6**
Item 01	Memory Strategies						
Item 02							
Item 03							
Item 04							
Item 05							
Item 06							
Item 07							
Item 08							
Item 09							
Item 10	Cognitive Strategies						.472
Item 11							
Item 12							
Item 13							
Item 14				.504			
Item 15							
Item 16				.466			
Item 17				.584			
Item 18							
Item 19							
Item 20							.443
Item 21							.404
Item 22					−.517		
Item 23							

Item 24	Compensation Strategies				−.587		
Item 25					−.504		
Item 26							
Item 27							
Item 28							
Item 29					−.588		
Item 30	Metacognitive Strategies						
Item 31							
Item 32							
Item 33		.573					
Item 34		.414					
Item 35		.644					
Item 36		.747					
Item 37		.752					
Item 38		.644					
Item 39	Affective Strategies					.598	
Item 40						.494	
Item 41						.544	
Item 42							
Item 43				.437			
Item 44							
Item 45	Social Strategies				−.457		
Item 46			−.536				
Item 47			−.579				
Item 48			−.884				
Item 49			−.626				
Item 50							

Note

1. The second column indicates the construct which the items are hypothesized to measure by the author (Oxford, 1990). Values below .40 are suppressed in this table. Rows which are shaded indicate items which did not appear on any factor above a value of .40.
2. Labels assigned to each factor were as follows: Factor 1, Metacognitive Strategies (Revised 6F); Factor 2, Social/Communicative Strategies (Revised 6F); Factor 3, Authentic Language Use Strategies (Revised 6F); Factor 4, Tolerance for Communicative Imprecision (Revised 6F); Factor 5, Affective Strategies (Revised 6F); Factor 6, Cognitive Strategies (Residual Items 6F).

As with the five-factor solution, half of the items in the SILL did not appear in the solution. Also, it is notable that, overall, the Affective Strategies items reappear in this solution (Factor 5, see below) and the Memory Strategies items—of which three items in the five-factor solution had appeared along with some Cognitive Strategies items in Factor 2—now completely disappear from the solution.

Factor 1 included precisely the same items (Items 33 through 38) as those which appeared on Factor 1 in the five-factor solution above. Factor coefficients for these items were all slightly lower

than they were for the five-factor solution; which was not unexpected because there is an additional factor in the model to take up variance. This factor was therefore labelled *Metacognitive Strategies (Revised 6F)*[2]. Comparative coherence across the five-factor and six-factor solutions was therefore maintained for these six items, because in both solutions they appeared together as a factor, and indeed, as the first factor in the respective solutions.

Factor 2 comprised precisely the same items (Items 46 through 49) as those which appeared on Factor 5 of the five-factor solution. This factor was therefore labelled *Social/Communicative Strategies (Revised 6F)*. The factor coefficients, but for one, were slightly higher than for Factor 5 in the five-factor model, and this is explained by this factor emerging earlier in this six-factor solution (Factor 2) than for the five-factor solution (Factor 5). Again, comparative coherence is maintained across the five-factor and six-factor solutions for these items, given that the same items appear together as a factor, albeit the factor position shifts in the order of factors extracted.

Factor 3 comprised the same three items (Items 14, 16 and 17) which had appeared on Factor 3 of the five factor solution (with very similar factor coefficients) plus one other item from the original Affective Strategies construct (Item 43). The content of this item was also, arguably, commensurate with the notion of authentic language use and therefore the label *Authentic Language Use Strategies (Revised 6F)* was also applied to Factor 3 in the six-factor solution. Here again, there is relative stability, or coherence, across solutions, other than for the appearance of one extra item (Item 43) along with the others in the six-factor solution, and this item had the weakest coefficient compared to the other three. Of course, while there is stability or coherence here across solutions, Oxford (1990) did not have an authentic language use construct in her taxonomy or in the SILL which expresses the taxonomy. It would seem on the basis of the evidence presented in these scores that *Authentic Language Use Strategies (Revised 5F)* and *Authentic Language Use Strategies (Revised 6F)* represent hitherto unspecified latents, or unanalyzed variance, lurking in scores produced by the SILL. Thus we can summarize this factor by saying that coherence is maintained across the two solutions for a newly identified latent which has not been previously specified by the author or previous researchers.

Factor 4 of this six-factor solution included all the same items which were included in Factor 4 of the five-factor solution, namely, one item from the author's Cognitive Strategies scale (Item 22), three items from the author's Compensation Strategies scale (Items 24, 25 and 29), and one item from the author's Social Strategies scale (Item 45). The respective factor coefficients for each of these items on each solution were also very similar. The label *Tolerance for Communicative Imprecision (Revised 6F)* was therefore assigned to Factor 4 on the six-factor solution, and coherence was maintained across the two solutions in this case. As with comments made with respect to *Authentic Language Use Strategies (Revised 5F)* and *Authentic Language Use Strategies (Revised 6F)*, immediately above, while coherence is maintained across solutions, the coherence is not with respect to a latent hypothesized by Oxford (1990), but rather with respect to a latent which has not previously been identified.

---

[2] As with the five-factor solution reported earlier, the term "Revised" in brackets indicates that while the name is the same as the original, the items comprising it are not. The "6F" indicates that the factor is revised within the six-factor solution, rather than within the five-factor solution reported above.

Factor 5 in the six-factor solution was new when compared with factors in the five-factor solution and comprised items from the original Affective Strategies construct (Items 39, 40 and 41). This was the construct which failed to gain representation in the five-factor solution, and effectively disappeared. Under the six-factor solution, it is rescued, but only with a relatively weak signal. Three of the original Affective Strategies items (Items 42, 43 and 44) do not appear in the factor, with two of these (Items 42 and 44) not appearing on any factor in the solution at all, and one (Item 43) appearing in the immediately-above reported *Authentic Language Use Strategies (Revised 6F)*. Factor 5 was therefore labelled *Affective Strategies (Revised 6F)*.

Factor 6 in the six-factor solution was weak with only three items (Items 10, 20 and 21) included, and these were only just so, because the factor coefficients were all just over .40 (the threshold for inclusion in the factor). No item had a coefficient above .50. Two of these items (20 and 21) had appeared under Factor 2, labelled *Cognitive Strategies (Revised 5F)*, in the five-factor solution. However, given that Factor 2 in the five-factor solution had included six items, it is difficult to claim that Items 20 and 21 represent the same sort of factor. This factor was therefore labelled *Cognitive Strategies (Residual Items 6F)*. The factor was qualified as "Residual Items" rather than "Revised," because it was too weak to claim a factor with reasonable operational expression. Although all three items did come from the author's original Cognitive Strategies scale, this scale has the largest number of items in the instrument (13 items). If 10 items have disappeared out of 13, this effectively means that 77 percent of the operational content of the scale is lost in the six-factor solution.

Discussion

When discussing the above results from a broad perspective, the most striking result for both solutions is that only 25 items, or half of the original items, appear in either solution; that is, either the five-factor or six-factor solution. Of these 25 items, 20 are common to both solutions including Items 14, 16, 17, 20, 21, 22, 24, 25, 29, 33, 34, 35, 36, 37, 38, 45, 46, 47, 48, 49. Under this analysis, it is clearly arguable that at least half of the SILL does not comprise any form of psychometric signal. Put another way, over half of the items produce scores which amount to noise. This overall result is consistent with recent research in the Greek population, for example, where Petrogiannis and Gavriilidou (2015) adopted EFA procedures which lead to a reduction in the length of the SILL to just 29 items.

On the issue of the number of factors which actually underlie the SILL, it is noteworthy that as far back as 18 years ago, and also in the Japanese population, Robson and Midorikawa (2001) reported initial solutions (in their Time 1 and Time 2 administrations of the instrument directed toward a research design for exploring test-retest reliability) calling for more than ten factors under the eigenvalue-greater-than-one rule. The PCA for the Time 1 and Time 2 administrations called for a 15-factor extraction and 13-factor extraction, respectively. Their decision, beyond seeking solutions with the aforementioned number of 15 and 13 factors (they did not use alternative means like parallel analysis to empirically obtain the number of factors to extract), was to also seek a six-factor solution in line with what the author has hypothesized. It is difficult, however, to compare the results in this study with the results for the Robson and Midorikawa

study, because their study is somewhat underreported in terms of the decision sequence for the EFA, and whether or not, for example, they sought rotated or un-rotated solutions. It does appear, by inference, that they may have sought an un-rotated PCA solution of six factors, because the first factor took up most of the variance which is typical of un-rotated PCAs. If the solution is un-rotated then the structure matrix and pattern matrix are the same, whereas in this study, with a rotated solution, the two are not the same; and of course, as is made clear above, the pattern matrix is interpreted in this study.

Turning more directly to the findings in this study, the main issue which presents for attention is the identification of aspects of coherence which may exist among the half of items which are included in the respective five-factor and six-factor solutions. Executing both the author's six-factor solution and the empirically determined five-factor solution was helpful in this, because the five-factor solution, while empirically supported by the parallel analysis, would have the necessary implication of forcing one of the originally hypothesized factors out of the solution. The six-factor solution does not, as a matter of necessity, do this; though it is possible that previously hypothesized factors do anyway disappear, but this would be due to other properties of the data, rather than the stipulation of less than six factors. Of course, if coherence in a group of items, by virtue of appearing on a factor together, extends beyond the five factor solution, and is observed also in the six-factor solution, this strengthens the case for the identification of this coherence as a signal in the data which is worth analyzing; and this applies for either how such signal may confirm or disconfirm what the author claims is measured via the SILL.

In this regard, therefore, the first factor (Items 33 through 38) extracted in both solutions, namely, *Metacognitive Strategies (Revised 5F)* and *Metacognitive Strategies (Revised 6F)*, would have to be cited. This factor rescued more of the items originally hypothesized to measure it (proportionately two thirds) than any other factor, and this was across both solutions. Nonetheless, it has to be said that three items originally hypothesized to measure metacognitive strategies did not appear in the factor in either solution. Also, Items 34 through 37 evidence very high skew and low means, with Items 34 and 35 being particularly weak in this respect.

Conversely, it would have to be mentioned that the original Memory and Cognitive Strategies scales are problematic in this set of data. In the five-factor solution, the second factor extracted seemed to represent a weak convergence of these two scales with items from both represented (Items 03, 04 and 07 [Memory]; and Items 19, 20, and 21 [Cognitive]). In the five-factor solution, this converged scale was labelled *Cognitive Strategies (Revised 5F)* because some of the original Memory Strategy items could also be construed as Cognitive Strategy items. Nonetheless, six of the original Memory Strategy items did not appear in the factor, and eight of the original Cognitive Strategy items also did not appear (with one of these [Item 22] appearing on Factor 4, and three [Items 14, 16 and 17] on Factor 3). When six factors were extracted, however, even this weak and converged picture was not maintained. None of the original author-hypothesized Memory Strategies items appeared on any factor at all, and only three of the originally hypothesized Cognitive Strategies items appeared just above the threshold factor coefficient of .40 in Factor 6. This factor was labelled *Cognitive Strategies (Residual Items 6F)* in the six-factor solution, because it appears to represent the weak residue, or left-overs, of the original Cognitive Strategies scale, in an overall picture, across both the five- and six-factor solutions, of weakness and general loss of signal. The

data in this study, therefore, suggests that the Memory Strategies scale and Cognitive Strategies scale are not properly operationalized, or properly differentiated. Extending the support for this position is the fact that three items from the original Cognitive Strategies scale (Items 14, 16 and 17), and mentioned immediately above, appear on Factor 3 in both solutions. All three of these seemed to invoke content which represented authentic language use as a strategy for learning, and this implies that there may be a latent present in the data which was not at all hypothesized by the original author, and which is operationalized by items presumed by the author (Oxford, 1990) to measure Cognitive Strategies.

Another possible, and previously not hypothesized, latent present in the data relates to Factor 4 in both the five- and six-factor solutions, and labelled in this study as *Tolerance for Communicative Imprecision (Revised 5F)* and *Tolerance for Communicative Imprecision (Revised 6F)*, respectively. In both cases, the factor includes Items 22, 24, 25, 29 and 45, with the last item, Item 45, appearing quite weakly and perhaps anomalously. The first four items all seem to include context which reflects some sort of tolerance for ambiguity or imprecision when communicating in the second or foreign language. This tolerance may be quite important for language learning because communicating with a language one is still learning invites imprecision, quite naturally, and a tolerance for this problem may assist with willingness to communicate and anxiety, as well as all sorts of problems which could potentially slow learning. In the case of Factor 4, therefore, and for both solutions, it is arguable that there is a signal in the data, but it is not the signal hypothesized by the author (Oxford, 1990). From a critical point of view, it is also worth pointing out that the items appearing on Factor 4 in both solutions could also be interpreted as encompassing comprehension strategies, because all of them basically reference an attempt at comprehension in as much as they also represent a sacrifice of precision put in service to this attempt.

Returning to items which do provide some support for what the author (Oxford 1990) claims is measured by the SILL, Items 46 through 49, and labelled *Social/Communicative Strategies (Revised 5F)* and *Social/Communicative Strategies (Revised 6F)* in this study, do align with what was originally hypothesized. The label used in this study adds the term "Communicative" to the term "Social," but it is arguable that these two terms are merely two sides of the same coin. Four out of the six items, originally hypothesized to measure the Social Strategies scale by the author, do appear alongside each other on the same factor and in both the five- and six-factor solutions.

Finally, and with respect to aspects of coherence in the data, it should be mentioned that while three of the items (Items 39, 40 and 41), out of six, hypothesized by the author to measure Affective Strategies appeared in the six-factor solution as Factor 5, labelled *Affective Strategies (Revised 6F)*, none of the items appeared on any factor in the five-factor solution. Inspection of the three items indicates content which is aligned with feelings of fear in the first two items and success in the third, but the scale is nonetheless weakly represented in the six-factor solution and disappears in the five-factor solution. Therefore, it is arguable that that Affective Strategies are not a strong point of coherence or signal in the data.

In conclusion of the overall results discussed above, it is clear that while there are aspects of support for some of what Oxford (1990) claims is measured by the SILL, there are also quite consequential departures and deficits. Principal among these is the number of items which effectively disappear under EFA, and whether this be the six-factor solution hypothesized by

Oxford, or the slightly simpler five-factor solution led by the empirical findings of the parallel analysis conducted in this study. The instrument is quite long, and one would have to ask whether many of these items serve any purpose, other than to contribute to operational noise, so to speak. The Metacognitive Strategies scale and the Social Strategies scale, originally hypothesized by Oxford, find some qualified support in these results. The majority of items from these scales do appear in both solutions on a single factor. Items on the Memory Strategies scale and Cognitive Strategies scale do not perform well in this analysis, and converge around a weak and hard-to-explain factor in the five-factor solution, and in the case of the six-factor solution either disappear (Memory Strategies scale) or find very weak expression (Cognitive Strategies scale). One aspect of the weakness of the items on the Cognitive Strategies scale is that some of them appear to be representing authentic language use as a strategy, which is a strategy not hypothesized by Oxford as underlying scores generated by the SILL. In other words, there is probably unspecified systematic variance in scores generated by the SILL (i.e. unspecified by Oxfords measurement model for the SILL), and authentic language use as a strategy may nominalize what at least some of this unspecified variance represents. Similarly, there is evidence for another source of systematic variance in this study, and not specified in the measurement model for the SILL, which relate to the items covering content implying a tolerance for communicative imprecision as a strategy for comprehension; notwithstanding the conceptual issue of whether a tolerance can be considered a strategy. At least some of the items which Oxford hypothesizes to measure Compensation Strategies might fall under this alternative analysis. The findings of systematic variance in the data which has previously been unspecified with respect to the SILL is probably one of the more important findings of this study, because it indicates that parts of the SILL which may be measuring something, do not necessarily measure something aligned with what the author claims is measured. Also, and even more importantly, these sources of systematic variance represent strategies for language learning which have probably not been given sufficient representation in taxonomies of strategies to date. Can authentic language use and tolerance for ambiguity as a comprehension strategy enter a taxonomy of language learning strategies as bona fide strategies in their own right? These are important questions for future research, and their appearance in the data for this study, may amplify their omission in the measurement model hypothesized for the SILL and the taxonomy on which the SILL is premised.

References

Anam, S., & Stracke, E. (2016). Language learning strategies of Indonesian primary school students: In relation to self-efficacy beliefs. *System, 60*, 1-10.

Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning, 28*(1), 69-83.

Bialystok, E. (1981). The role of conscious strategies in second language proficiency. *The Modern Language Journal, 65*, 24-35.

Chamot, A. U., & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals, 22*, 13-24.

Chamot, A. U., & O'Malley, J. M. (1987). The cognitive approach: A bridge to the mainstream. *TESOL Quarterly, 21*, 227-249.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104.

El-Dib, M. A. B. (2004). Language learning strategies in Kuwait: Links to gender, language level, and culture in a hybrid context. *Foreign Language Annals, 37*, 85-95.

Green, J., & Oxford, R. L. (1995). A close look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly, 29*, 261-297.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of Coefficient Alpha as an Index of Test Unidimensionality. *Educational and Psychological Measurement, 37*, 827-838.

Hong-Nam, K., & Leavell, A. G. (2006). Language learning strategy use of ESL students in an intensive English learning context. *System, 34*(3), 399-415.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Isemonger, I. M. (2016). A confirmatory approach to the structural validity of scores generated by the Strategy Inventory for Language Learning (SILL). *International Journal of Social and Cultural Studies, 9*, 1-14.

McCullen, M. (2009). Using language learning strategies to improve the writing skills of Saudi EFL students: Will it really work? *System, 37*(3), 418-433.

Naiman, N., Frohlich, M., & Todesco, A. (1975). The good second language learner. *TESL Talk, 6*, 68-75.

Nisbet, D. L., Tindall, E. R., & Arroyo, A. A. (2005). Language learning strategies and English proficiency of Chinese university students. *Foreign Language Annals, 38*(1), 100-107.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396-402.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. P. (1985a). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*(1), 21-46.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. P. (1985b). Learning strategy applications with students of English as a foreign language. *TESOL Quarterly, 19*, 557-584.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston, MA.: Heinle and Heinle.

Petrogiannis, K., & Gavriilidou, Z. (2015). Strategy inventory for language learning: findings of a validation study in Greece. . In M. Carmo (Ed.), *Education Applications and Developments* (pp. 223-236). Lisboa: In Science Press.

Piaget, J. (2001a). *The language and thought of the child*. London: Routledge.

Piaget, J. (2001b). *The psychology of intelligence*. London: Routledge.

Robson, G., & Midorikawa, H. (2001). How reliable and valid is the Japanese version of the strategy inventory for language learning (SILL)? *JALT Journal, 23*(2), 202-226.

Rubin, J. (1975). What the 'good language learner' can teach us. *TESOL Quarterly, 9*, 41-51.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 2*, 117-131.

Stern, H. H. (1975). What can we learn from the good language learner? *The Canadian Modern Language Review, 31*, 304-318.

Sun, M., Mantero, M., & Summers, R. (2014). Chinese adult second language learners' learning strategy and communicative strategy use. *Journal of Second and Multiple Language Acquisition, 2*(1), 22-35.

Tam, C. H. (2013). A study on language learning strategies (LLSs) of university students in Hong Kong. *Taiwan Journal of Linguistics, 11*(2), 1-42.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1986). *Thought and language*. Cambridge, MA: MIT Press.

Watson, J. R., & Ebner, G. (2018). Language learning strategy use by learners of Arabic, Chinese, and Russian during study abroad. In C. Sanz & A. Morales-Front (Eds.), *The Routledge Handbook of Study Abroad*. New York: Routledge.

Wharton, G. (2000). Language learning strategy use of bilingual foreign language learners in Singapore. *Language Learning, 50*, 203-243.

Wilkinson, L., & APA Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*(8), 594-604.

Yang, N. D. (1999). The relationship between EFL learners' beliefs and learning strategy use. *System, 27*, 515-535.

Zwick, W. R., & Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research, 17*, 253-269.

Zwick, W. R., & Velicer, W. F. (1986). Factors influencing five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

日本語版言語学習戦略テストの潜在的構造：探索的因子分析を用いたデータ駆動型アプローチ

アイズマンガー・イアン

Strategy Inventory for Language Learning（Oxford, 1990）、または SILL は、語学学習者が使用する方略を測定すると主張されており、文献の中でのその存在は重要である。この測定法は、日本語を含め様々な言語に翻訳されている。これらの翻訳版の１つである Isemonger（2016）の日本版では、原著者であるオックスフォードの研究仮説を確証的因子分析（CFA）により心理測量が検証されたが、その仮説は支持されなかった。よって本研究では、探索的因子分析（EFA）を使用し、データへのポステリオリアプローチを使用することとした。この分析により、仮説モデルとの適合の欠如を考慮して生成されたスコアの潜在的構造がどのようなものであるかを探索した。主な結果は、項目の半分は検出された２つの EFA 因子解の中に検出されず、残りの項目はオックスフォードが仮定した尺度と部分的に一致していた。そのことから、これまでに原著者によって記述されていない潜在性がデータの中に存在する可能性があることが示唆された。

Keywords: Strategy Inventory for Language Learning, SILL, Learning Strategies, Exploratory Factor Analysis, Unspecified Latents