The Modern Language Journal

Strategy Inventory for Language Learning–ELL Student Form: Testing for Factorial Validity

YULIYA ARDASHEVA
Washington State University Tri-Cities
ESL/Bilingual Education
Richland, WA 99354
Email: jearda@gmail.com

THOMAS R. TRETTER
University of Louisville
Department of Middle and Secondary Education
Louisville, KY 40292
Email: tom.tretter@louisville.edu

As the school-aged English language learner (ELL) population continues to grow in the United States and other English-speaking countries, psychometrically sound instruments to measure their language learning strategies (LLS) become ever more critical. This study adapted and validated an adult-oriented measure of LLS (50-item *Strategy Inventory for Language Learning* [SILL]; Oxford, 1990) for school-aged ELLs in a sample of 1,057 elementary, middle, and high school students. The two-stage validation process resulted in a shorter, 28-item version of the instrument, which we entitled the *SILL–ELL Student Form*. The results of confirmatory factor analyses indicated a good fit to the validation (GFI = .92, CFI = .93, RMSEA = .04, SRMR = .05) and combined (GFI = .95, CFI = .95, RMSEA = .03, SRMR = .04) samples. The new measure has strong psychometric characteristics for use with school-aged ELLs to diagnose their use of LLS in six distinct categories and is approximately half as long as the original SILL, which enhances its pragmatic value for busy classrooms. Additionally, this study addressed some of the theoretical issues with strategy categorization noted in the literature. Other applications of the measure for practice and research are discussed.

LANGUAGE LEARNING STRATEGIES (LLS) have been defined as "specific actions consciously employed by the learner for the purpose of learning language" (Griffiths, 2007, p. 91). Oxford (1999) likewise described learning strategies as "specific actions taken by the learner to make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (p. 8). Higher strategy use has been associated with higher proficiency in a second language (L2; e.g., Ardasheva, 2011; Dreyer & Oxford, 1996; G. Hu et al., 2009; Lan & Oxford, 2003; Peacock & Ho, 2003; see also Cohen & Macaro, 2007), better academic performance in content areas such as language arts and mathematics (Ardasheva & Tretter, in press; Chamot et al., 1992; Montes, 2002; see also Chamot, 2007), and higher performance on cognitive/behavioral measures (e.g., self-efficacy: Magogwe & Oliver, 2007; motivation: MacIntyre & Noels, 1996; Nunan, 1997; Oxford & Nyikos, 1989; Schmidt & Watanabe, 2001).

There is evidence to suggest that LLS can be successfully taught (Graham & Macaro, 2008; Gunning, 2011; G. Hu et al., 2009; Nunan, 1997; Vandergrift & Tafaghodtari, 2010; see also Hassan et al., 2005; McDonough, 1999). Yet, empirical evidence, particularly with regard to the relationship between L2 learning and LLS, remains inconsistent (Gardner, Tremblay, & Masgoret, 1997; Nisbet, Tindall, & Arroyo, 2005) with some new evidence suggesting that this relationship may be moderated by age and length of L2 study (Ardasheva, 2011). Instrumentation that would accurately diagnose LLS, then, is key for

The Modern Language Journal, 97, 2, (2013)
DOI: 10.1111/j.1540-4781.2013.12011.x
0026-7902/13/474–489 $1.50/0

informing L2 learners' learning and their teachers' teaching practices, as well as a desirable research tool for clarifying inconsistencies in previous research.

Currently, the most frequently used instrument in LLS research is Oxford's (1990) *Strategy Inventory for Language Learning* (SILL). The SILL groups LLS into six categories: (a) *memory strategies:* used for information storage and retrieval; (b) *cognitive strategies:* used for comprehension and production; (c) *compensation strategies:* used to overcome limitations in linguistic knowledge or performance; (d) *metacognitive strategies:* used to plan, organize, focus, and monitor learning; (e) *affective strategies:* used to control motivation and emotions; and (f) *social strategies:* used for cooperative interaction with others. This adult-oriented instrument has been translated into 17 languages and appears in dozens of major publications, as well as in over 40 dissertations (Oxford, 1999), involving the study of LLS among second and foreign language learners.

However, the structural validity of the SILL is far from well established (Hsiao, 1997; Hsiao & Oxford, 2002), particularly for school-aged English language learners (ELLs; Oxford & Burry–Stock, 1995), a rapidly growing school population in the United States and other English-speaking countries (Goldenberg, 2008; Kaufman & Crandall, 2005; Kim & Jang, 2009). In a review, Oxford and Burry–Stock (1995) reported that among six SILL validation studies, only one focused on school-aged L2 learners and only two on ELLs. The primary purpose of the current study was to adapt, via a systematic, empirical process, the theoretical, six-category model of the SILL instrument (Oxford, 1990) for school-aged ELLs and then to test the modified instrument with this important student population.

CONCEPTUAL FRAMEWORK

Defining Language Learning Strategies

Learner-strategy researchers have adopted their definitions of learning from psychology, where learning is commonly referred to as the process of storing and retrieving information (Dörnyei, 2005; Rubin, 1981). Strategies have been described as techniques or devices learners use to gain knowledge (Rubin, 1975) or as actions or steps toward achieving a given objective (Cohen, Weaver, & Li, 1996; Oxford, 1990). In defining LLS, several authors emphasize the role of consciousness (Cohen, 1998; Cohen et al., 1996; Macaro, 2006). Cohen (1998) argues that "the element of consciousness is what distinguishes *strategies* from those processes that are not strategic" (p. 4, emphasis in original). He notes that language learners, especially older learners, may have a keen awareness of what language features should be learned. In a similar vein, Genesee et al. (2005) observed that the use of LLS is characteristic of L2 development because learners typically begin to learn a new language at a more mature age and, thus, can draw on conscious, explicit strategies to enhance their learning.

LLS and Autonomous Learning

Although currently no single, commonly accepted theoretical framework concerning LLS exists[1] (Macaro, 2006; Rivera–Mills & Plonsky, 2007; Skehan, 1998; for an excellent bibliography on LLS research and theory development, see Oxford, 2011), two cognitive and social theories of learning—namely, Anderson's Adoptive Control of Thought (1982, 1989) and Vygotsky's (1978) sociocultural theory—have been particularly informative in guiding LLS research and in illuminating findings (Chamot & O'Malley, 1994; Macaro, 2006; O'Malley et al., 1985; O'Malley, Chamot, & Walker, 1987). Building on these theories, many researchers and educators (e.g., McDonough 1999; Oxford, 1999; see also Cohen & Macaro, 2007) argue that LLS contribute to *learner autonomy*, the ability and willingness to take responsibility for one's own learning.

O'Malley et al. (1985), for example, argue that if learning is generated through cognitive processing, then, "strategies that promote the greatest amount of mental activity should result in the most learning" (p. 24). McDonough (1999) notes that LLS researchers commonly perceive strategies as learning-to-learn activities and claims that strategies promote self-regulated learning by facilitating student progression through the stages of skill acquisition—cognition, association, and autonomy—as theorized by Anderson (1982, 1989). Macaro (2006) has gone even further in arguing that strategies do not simply make learning more independent and efficient, but are "the raw material without which L2 learning cannot take place" (p. 332).

Similar to McDonough, Oxford (1999) suggests that LLS are key to the learner's autonomous learning. According to Oxford, strategies are learner-initiated actions directed at improving learning; thus, by definition, strategies "reflect the

learners' degree of autonomy and are mechanisms by which the learner develops still greater autonomy" (p. 111). While recognizing learner autonomy as an individual characteristic, Oxford also stresses the importance of social factors in developing student self-regulated learning behaviors. Drawing on Vygotsky's (1978) ideas, Oxford observes that strategies are internalized via social interaction; thus, the development of an independent, problem-solving learner also relies on the help of more capable others, including peers and teachers. In a similar vein, McDonough (1999) suggests that, while some language learners deploy LLS spontaneously, others need instruction to enact the strategies independently.

Strategy Categorization Schemas

Early LLS research (e.g., Rubin, 1975, 1981) identified and described a rather large number of strategies learners employ. While discrepancies in defining some specific strategies and in describing hierarchical relations among LLS remain (Griffiths, 2007; O'Malley et al., 1985), and the concept of learning strategies itself has met some criticism in the literature (Dörnyei, 2005), several strategy categorization schemas have been proposed and applied in both research and language classroom settings.

According to Hassan et al. (2005), most existing strategy schemas include at least three categories: *metacognitive, cognitive*, and *socioaffective*. This three-category domain, in particular, has been endorsed by O'Malley et al. (1985) and Chamot and O'Malley (1994). Cohen et al. (1996), however, further distinguish between *language learning strategies*, defined as conscious behavior aimed at facilitating language learning, and *language use strategies*, defined as behavior that facilitates communication in a new language. Among language learning strategies, Cohen et al. identify four types: (a) *cognitive:* strategies facilitating perception, storage, and retrieval of linguistic elements, (b) *metacognitive:* planning, organization, monitoring, and evaluation strategies, (c) *social:* strategies facilitating cooperation and feedback, and (d) *affective:* strategies regulating motivation, emotions, and attitudes. Among language use strategies Cohen et al. distinguish between *performance strategies*, such as rehearsal, that facilitate execution, and *communicative strategies*, such as circumlocution, that help get the message across despite linguistic gaps.[2]

Unlike Cohen and his colleagues, Rubin (1981) and Oxford (1990) do not draw an explicit distinction between strategies for language learning and those for language use. Instead, they distinguish between *direct* and *indirect* strategies. Oxford considers *direct* strategies to be those that require mental processing of linguistic information and *indirect* strategies as those that "support and manage language learning without (in many instances) directly involving the target language" (p. 135). Similarly, Rubin suggests that *direct* strategies (clarification/verification, monitoring, memorization, guessing/inductive inference, deductive reasoning, and practice) contribute directly to learning whereas *indirect* strategies (creating opportunities for practice, using production tricks) permit learning. Oxford's (1990) LLS taxonomy, in particular, details three direct and three indirect strategies.

Among direct strategies, Oxford (1990) includes *memory, cognitive*, and *compensation*. *Memory* strategies (e.g., grouping, using context) help the learner to store and retrieve information, to cope with a large L2 vocabulary, as well as to move from factual (knowledge) to procedural (skills) levels through automatization. *Cognitive* strategies (e.g., practicing, skimming, and note-taking) involve transformation of the new linguistic material to aid comprehension and production. *Compensation* strategies (e.g., guessing, asking for help, and using gestures) enable students to use language despite gaps in linguistic knowledge and, thus, provide for more practice. Among indirect strategies, Oxford includes *metacognitive, affective*, and *social*. *Metacognitive* strategies (e.g., focusing attention, planning for learning, and evaluation) enable learners to control their cognition. *Affective* strategies (e.g., lowering anxiety and self-encouragement) assist in regulating emotions, attitudes, and motivation. Finally, *social* strategies (e.g., questions and cooperation) facilitate interaction.

Rationale for Selecting the SILL Categorization Schema

The SILL, the instrument at the center of this study, is grounded in Oxford's (1990) six-factor LLS taxonomy and was chosen for three reasons. First, a recent confirmatory factor analysis (CFA; Hsiao & Oxford, 2002), which investigated a total of 15 competing models[3] based on strategy categorization schemas discussed in the previous section, established that Oxford's six-factor taxonomy provided the most consistent account for student data.

Second, the SILL remains "the most widely used strategy questionnaire in the world"

(Chamot, 2004, as cited in Gunning, 2011, p. 71). Further, while a recent and highly promising trend in strategy research is to use task-based strategy measures (e.g., G. Hu et al., 2009; Tseng, Dörnyei, & Schmitt, 2006; Vandergrift et al., 2006; see also discussions in Dörnyei, 2005; Hsiao & Oxford, 2002; Oxford et al., 2004) because they facilitate a more focused investigation of the relationships between strategies and specific language domains (e.g., speaking) or skills (e.g., grammar), the general SILL has a broader range of potential applications. Examples of such applications include investigations of the relationships between strategy use and student cognitive/behavioral and academic outcomes (e.g., Ardasheva & Tretter, in press; MacIntyre & Noels, 1996; Magogwe & Oliver, 2007; Oxford & Nyikos, 1989). Further, as Hsiao and Oxford (2002) observe, the general SILL is less demanding in terms of administration and has yielded research findings comparable to those produced by research using task-based measures.

Finally, Oxford's (1990) scheme offers certain advantages for classrooms, both language and content, in that it foregrounds a unique set of strategies that are highly appropriate for the instructional settings and learning goals at hand (Chamot, 2009). Previous research has revealed positive associations between memory strategies and vocabulary knowledge (Takeuchi, 1993), which suggests the appropriateness of such strategies for vocabulary development. Metacognitive and cognitive strategies, assessed by the SILL or other LLS measures, in turn, have been linked to a broader range of language outcomes (listening: Peacock & Ho, 2003; Takeuchi, 1993; Vandergrift et al., 2006; structures: Peacock & Ho, 2003; Takeuchi, 1993; reading: Clarke, 1979, 1980; Peacock & Ho, 2003; Schoonen, Hulstijn, & Bossers, 1998; speaking and writing: Peacock & Ho, 2003; Takeuchi, 1993) and academic outcomes in content areas (Chamot et al., 1992; Montes, 2002). Social and affective strategies, Chamot (2009) observes, would be helpful when students are focused on the development of social/conversational language skills (Cummins, 2008).

Research on the Validity of the SILL

In their review of the SILL validity research, Oxford and Burry–Stock (1995) reported the results of exploratory factor analyses (EFA) that examined the underlying structure of the instrument using data sets from six studies, only one of which focused on school-aged L2 learners and only two on ELLs. They noted strong evidence of the instrument's content validity (i.e., an examination of how well SILL items related to a comprehensive LLS taxonomy by a panel of experts), predictive validity (i.e., findings from studies relating LLS to L2 proficiency),[4] and concurrent validity (i.e., findings from one study indicated that LLS preference and use varied with learning sensory preferences). With regard to the construct validity of the SILL, findings were less conclusive.[5] Further, although a recent CFA investigation (Hsiao & Oxford, 2002; discussed in the previous section) established that, among fourteen competing LLS taxonomies examined, Oxford's six-factor taxonomy provided the most consistent account of college student data, the fit indices indicated that the model did not offer an adequate fit to the data,[6] which led the authors to conclude that there was still substantial room for instrument improvement. Among possible SILL improvements, Hsiao and Oxford mentioned the following: (a) distinguishing items that are appropriate for second versus foreign language contexts (i.e., while watching shows in L2 context represents a conscious learning strategy on the part of a foreign language learner, the same behavior may simply represent an everyday reality for a second language learner); (b) reclassifying strategies to produce a clearer distinction among categories (i.e., rephrasing or eliminating strategies that may be categorized in more than one category); and (c) establishing the same level of item specificity (i.e., ensuring that items are non-generic and worded in a way that clarifies the context of strategy application for all respondents).

SILL Research With School-Aged L2 Learners

Several studies (e.g., Chen, 2009; Gunning, 1999, 2011; Kaylani, 1996; Lan & Oxford, 2003; Magogwe & Oliver, 2007) used the SILL to examine strategy use among school-aged English-as-a-second/foreign-language learners. The overall result was that more successful students used more strategies than did less successful students (Kaylani, 1996; Lan & Oxford, 2003; Magogwe & Oliver, 2007). Notably, while children used the same strategy categories as adolescents and adults, there was a difference in preference. Elementary school students preferred affective, compensation (Gunning, 1999, 2011), and social (Magogwe & Oliver, 2007) strategies. Middle school students reported greater use of social,

metacognitive, affective, memory, and cognitive strategies; high-school-aged students indicated a strong preference for compensation (Chen, 2009) and metacognitive (Magogwe & Oliver, 2007) strategies. Magogwe and Oliver (2007) found a strong association between overall strategy use and ELL student self-efficacy beliefs.[7] All of the abovementioned studies reported simplifying, translating, or shortening the SILL for use with school-aged L2 learners. Although the authors presented reliability coefficients for the modified SILL (Chen, 2009; Gunning, 1999, 2011; Magogwe & Oliver, 2007) and/or established its content validity (i.e., content analysis by experts: Gunning, 1999, 2011; back translation: Gunning, 1999; piloting: Gunning, 1999, 2011; Magogwe & Oliver, 2007), none of the studies involved a thorough examination of the psychometric properties of the modified versions of the SILL. The primary purpose of this study was to modify and validate the SILL (Oxford, 1990) for school-aged ELL students.

METHOD

Study Design

Recognizing (a) the limited nature of empirical evidence for either supporting or refuting the adequacy of the 50-item SILL for school-aged ELLs (Oxford & Burry–Stock, 1995), (b) the common practice of simplifying the SILL for younger student populations (e.g., Gunning, 1999, 2011), and (c) the existing criticisms and recommendations for enhancing the instrument's validity (Hsiao & Oxford, 2002), the present study had two stages: instrument calibration (Stage 1) and instrument validation (Stage 2).

Stage 1 began with the initial instrument modifications (i.e., simplifying the wording of the items) and the evaluation of the fit of the fifty-item SILL model in Sample A using CFA techniques. This analysis had two main purposes: to establish a baseline (comparison) assessment of the model fit to ELL student data and to draw on structural equation modeling techniques to inform further instrument modifications by identifying potentially problematic items. Next, guided by the language learning strategy literature (e.g., Cohen, 1998; Hsiao & Oxford, 2002), the results of CFA and EFA analyses, and input from a panel of experts (an early childhood educator, an elementary teacher with background in ESL, and an ESL teacher) a shorter and more appropriate version of SILL for use with school-aged ELLs was developed. To reflect the measure's new intended audiences, the modified instrument was entitled *SILL–ELL Student Form*.

Stage 2 involved two steps. First, the structural validity of the modified instrument was tested with CFA techniques in Sample B. Second, the psychometric properties of the instrument were further evaluated in the Combined Sample.

Sample and Procedure

This study was an analysis of data collected for a research project involving school-aged ELL students receiving English-as-a-second-language (ESL) services and enrolled in grade-level classrooms (Grades 3–8, 10–11) in a large Midwestern urban school district in the United States. The sample included 1,057 ELL students (651 elementary, 275 middle, and 131 high school; 48% female, $M_{age} = 12.21$, $SD = 2.80$, age range: 9–21 years) attending 38 schools. The students spoke over 40 native languages. The most frequently spoken languages were Spanish (48.3%), Maymay (12.5%), Somali (7.2%), Karen (3.8%), and Turkish (3.5%). Approximately 93% of students were eligible for free- or reduced-price lunches. Average time in U.S. schools was 42.87 months ($SD = 21.44$), or about 3.6 years. The SILL questionnaire was administered during regular instructional time by ESL teachers at the end of the school year. The teachers were instructed to read and explain the directions to the students and to provide language accommodations that had been used with the students as part of regular instruction during the school year.

Instrumentation

The Strategy Inventory for Language Learning (SILL version 7.0 for ESL/EFL learners; Oxford, 1990) is a self-report instrument designed to capture strategy use by language learners. This adult-oriented instrument comprises 50 items grouped into six categories: (a) memory strategies (9 items); (b) cognitive strategies (14 items); (c) compensation strategies (6 items); (d) metacognitive strategies (9 items); (e) affective strategies (6 items); and (f) social strategies (6 items). The answer categories are structured on a five-point scale ranging from 1 = *never or almost never true of me*, to 5 = *always or almost always true of me*. The average reliability coefficients (Cronbach's alphas) reported in studies ranged from .67 (Hong–Nam & Leavell, 2006) to .95 (Dreyer & Oxford, 1996).[8]

To ensure comprehension by elementary ELL students (the youngest students in the study), the wording of some original SILL items was simplified by a panel of elementary and ESL education specialists, including an early childhood educator, an elementary teacher with background in ESL, and an ESL teacher. After modifications, the readability level of the instrument was 3.2. Following Dillman's (2007) recommendations, the simplified SILL was piloted with a group of 6 third-grade ELLs with varied levels of English proficiency. Pilot testing followed a think-aloud format (i.e., the respondents read or listened to the questions and verbalized their thinking). The results of the pilot indicated that the simplified items were appropriate for elementary ELL students; in most cases, students correctly understood both the items' content and intent. The problematic items were further modified based on student feedback (e.g., children understood the word "student" more easily than the word "learner"). Simplified SILL items are displayed in Appendices A and B.

DATA ANALYSES

Instrument validation procedures proceeded with a calibration–validation design, with the total sample randomly split into Sample A ($n = 529$) and Sample B ($n = 528$) supported by baseline CFA, a panel of experts, and EFA in Sample A (Stage 1); CFA in Sample B and further examination of the psychometric properties of the instrument in the Combined Sample (Stage 2). All CFA analyses involved the use of AMOS software (Arbuckle, 2008).

CFA Analytical Approach

Because structural equation modeling CFA applications specify an a priori theoretical model, the assessment of how well CFA models fit the data was of primary concern (Byrne, 2001, 2010). In the case of strictly confirmatory structural equation modeling applications,[9] "the model is a given at the beginning of the analysis, and one of the main questions to be answered is whether it is supported by the data" (Kline, 2005, p.10). The assessment of fit of theoretical (single-solution) models typically involves the evaluation of the χ^2 goodness-of-fit statistic, supplemented by a set of more easily interpretable fit indices that quantify "the degree of fit along a continuum" (L. Hu & Bentler, 1999, p. 2). Because individual fit indices have limitations (e.g., sensitivity to sample size or model complexity), the structural equation modeling literature (e.g., L. Hu & Bentler, 1999; Marsh & Hau, 1996) recommends using combinations of fit indices from different measurement families.

The specific indices assessed in this study included: goodness-of-fit index (GFI), comparative fit index (CFI), the standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA). The first two indices represent goodness-of-fit and incremental fit index families, respectively; the latter two are residual-based. As indicators of acceptable model fit, this study used values close to .95 for GFI (Shevlin & Miles, 1998) and CFI (Bentler, 1990) and values close to .08 and .06 for SRMR and RMSEA (L. Hu & Bentler, 1999), respectively. Finally, theoretically meaningful modification indices (MIs; "the expected drop in overall χ^2 value if the parameters were to be freely estimated in a subsequent run," Byrne, 2001, p. 90) served to identify potential areas of model misfit due to item cross-loadings and error correlations. According to Byrne (2001), while the presence of cross-loadings suggests that an item may capture more than one construct, error correlations represent systematic measurement error that may arise from item characteristics[10] and may represent either a small omitted factor or an overlap in the items' content.

Data Preparation

The original data set was screened for unusual patterns and missing data. The results indicated that approximately 26% of the SILL surveys had missing data on at least one item. (This is not unexpected in light of the 50-item length of the instrument.) The item with the largest amount of missing data had only 2.5% (i.e., 26 out of 1,057) missing scores. Thus, because no individual item had a large amount of missing data, the missing scores were imputed using SPSS's series means function (i.e., replacing the missing value with the sample-based item mean score). All kurtosis and skewness values for individual items fell in the range of -1.4 to $+1.4$ and were judged as being acceptable indicators of data normality.

RESULTS

Stage 1

Stage 1 involved two steps. First, to establish a baseline assessment of the model fit and to identify potentially problematic items, the factorial validity of the 50-item SILL was evaluated in

Sample A using a CFA approach. Second, guided by the language learning strategy literature (e.g., Cohen, 1998; Hsiao & Oxford, 2002), the results of CFA and EFA analyses, and input from a panel of experts, a shorter and more appropriate version of SILL for use with school-aged ELLs was developed.

50-Item SILL: Model Testing in Sample A. To establish a baseline, the authors first estimated the SILL theoretical model on the basis of Oxford's (1990) six-factor strategy taxonomy and assessed the model fit to Sample A data. The Hoelter statistic of 283 ($p < .01$) was larger than its recommended cut-off criterion of 200 (Byrne, 2001); this indicated that the sample size was adequate for testing the model. The theoretical model specified four a priori hypotheses: (a) six factors would adequately explain the school-aged ELLs' responses to SILL; (b) each SILL item would load on the SILL category it was designed to measure, and have a zero loading on all other SILL categories; (c) the six SILL categories would correlate; and (d) measurement error terms would not correlate. The overall hypothesis was that the SILL theoretical model would adequately fit the variance/covariance structure in the school-aged ELLs' response data.

The results of the theoretical model analyses with Sample A indicated, at best, a moderate fit to the school-aged ELL data (see Table 1). Both GFI of .82 and CFI of .85 (an absolute fit index and a comparative fit index, respectively) were smaller than their recommended cut-off criterion of close to .95. This indicated that the model did not adequately explain student data. In contrast, the small values of residual-based absolute fit indices, RMSEA = .045, CI 90% [.042; .047], and SRMR = .054, indicated a close fit between the model-specified variance/covariance matrix and the population variance/covariance matrix. Both statistics, however, have limitations and provide only an overall estimation of the model fit: The SRMR depends on sample size (i.e., as sample size increases, SRMR decreases) and the RMSEA depends on the complexity of the model (i.e., as the number of parameters increases, RMSEA decreases). The examination of the residual matrix, a measure of how well the relationships among individual model items are specified (Raykov & Marcoulides, 2006), identified four items with five or more large, statistically significant residual values (i.e., > 2.58); this indicated the presence of underexplained relationships among observed variables under two or more strategy categories. The overall results indicated that one or more of the four a priori hypotheses did not adequately define the model.

Areas of Misfit. To identify possible areas of misfit (i.e., items accounting for substantial proportions of model misspecification), we examined MIs associated with regression weights and error covariances: The results indicated the presence of both item cross-loadings and error correlations.

The total number of statistically significant MIs associated with regression weights was 10 ($p < .05$). For example, the item "I don't let my English mistakes stop me from speaking" cross-loaded on four factors. Because this item could be theoretically categorized as a strategy that is affective (controlling emotions), metacognitive (thinking about thinking), cognitive (increasing opportunities for practice), and social (increasing opportunities for interaction), the item may represent a case of category overlap as has been noted in the literature (e.g., Cohen, 1998; Hsiao & Oxford, 2002).

The total number of statistically significant MIs associated with error covariances was 63 ($p < .001$). The larger MI associated with error covariances, for example, was that between the following two items: "I often start talking in English with others" and "I watch TV shows and movies in English." These two items may represent an omitted factor that is pertinent to the context of the study rather than to cognitive

TABLE 1
Goodness-of-Fit Statistics Across SILL Models

Model	χ^2	*df*	GFI	CFI	RMSEA	90%CI	SRMR
50-item SILL							
Sample A ($n = 529$)	2384.12	1160	.82	.85	.045	[.042, .047]	.054
28-item SILL							
Sample B ($n = 528$)	582.81	335	.92	.93	.037	[.032, .042]	.048
Total Sample ($N = 1{,}057$)	738.19	335	.95	.95	.034	[.030, .037]	.038

Note. GFI = goodness-of-fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation; 90%CI = 90% confidence interval; SRMR = standardized root mean square residual.

strategy use. In other words, if, in foreign language contexts, seeking opportunities to enhance exposure to an L2 (either through communication or through exposure to L2 media) may represent conscious learning behaviors, in second language environments, these actions more likely reflect the everyday reality of the L2 learners rather than their learning behaviors. Such item appropriateness for second versus foreign language contexts has been noted in LLS research (e.g., Hsiao & Oxford, 2002; Oxford & Burry–Stock, 1995). Further, examining the number of statistically significant MIs associated with individual items identified 10 items with a potential content overlap with five to eight other items. These empirical findings, supported by theoretical considerations and input from a panel of experts (described in the following section), guided the development of the SILL–ELL Student Form.

The Development of the SILL–ELL Student Form. As suggested in critiques of the instrument noted in the literature, the inadequate fit of the 50-item model may be attributable to the instrument itself. First, many of the SILL items are worded in general abstract terms (Cohen, 1998; Hsiao & Oxford, 2002), which may be less suitable for younger learners. Second, the original instrument was developed for use in both second and foreign language contexts (Oxford, 1990); thus, some items may be irrelevant to one or the other learner group (Hsiao & Oxford, 2002). Third, certain items can theoretically fit more than one strategy category which results in substantial category overlap (Cohen, 1998; Hsiao & Oxford, 2002; Macaro, 2006). The latter critiques are consistent with the CFA results, which indicated that a number of SILL items cross-loaded on two to four strategy categories and provided evidence of either item overlap or an omitted factor.

On the basis of the LLS literature described in the previous section, a more appropriate version of the SILL for use with school-aged ELLs was developed next. The procedure involved three steps. In Step 1, two specialists in early childhood and ESL rated the 50-item SILL on item-specificity on a 3-point scale ranging from 1 = *concrete* (single possible interpretation) to 3 = *abstract* (multiple possible interpretations). In Step 2, two ESL specialists rated the 50-item SILL as either appropriate or not appropriate for L2 learning contexts on a 2-point scale. When rating items as "non-appropriate," the panelists were to provide a rationale. Next, items (a) rated by the panelists as generic or non-appropriate for L2 contexts and (b) showing evidence of measurement error (as suggested by the CFA results reported earlier) were deleted from the instrument. The number of items eliminated at this stage was 14.

In Step 3, EFA in Sample A served to identify and eliminate strategies that might be categorized in more than one category. The Kaiser–Meyer–Olkin statistic of .95 indicated that the sample size was adequate for the procedure. Principal component analysis and varimax rotation with Kaiser normalization served as the extraction and rotation methods, respectively. Factor extraction decisions were made on the basis of eigenvalues greater than one and the leveling-off point on the scree plot. A six-factor solution best described student data. A factor loading of .387 (i.e., at least 15% of the shared variance between the item and the corresponding factor loading; Stevens, 2002) served as the decisional criterion for statistical and practical significance testing. On the basis of the EFA results, corroborated by CFA results reported earlier, 28 original SILL items that showed no evidence of cross-loading, that loaded on categories they were designed to measure, and that were consistent with requirements of factor interpretability were retained. (Appendix A displays descriptive statistics for all retained items; items deleted by the described procedures appear in Appendix B.) In Stage 2, the structural validity of the 28-Item SILL–ELL Student Form was examined by using CFA techniques.

Stage 2

Stage 2 had two steps. First, the structural validity of the modified instrument was tested with CFA techniques in Sample B using the cross-validation approach. Second, the psychometric properties of the SILL–ELL Student Form were further evaluated in the Combined Sample.

28-Item SILL: Model Validation in Sample B. The results of the 28-Item SILL–ELL Student Form model analyses in Sample B indicated an acceptable fit to the school-aged ELL data, GFI = .92, CFI = .93, RMSEA = .037, CI 90% [.032; .042], and SRMR = .048 (see Table 1). They suggested that the model tested could be accepted without further modification in view of its adequate correspondence to the school-aged ELL student data. These results also indicated that the individual items under each strategy category tapped into the same domain with an adequate degree of certainty. The new instrument

preserved the original SILL six-factor structure and represented a substantial improvement in model fit compared with the 50-item SILL (see Table 1).

Combined Sample Analyses. The final CFA model evaluated in this study was fitted to the Combined Sample ($N = 1{,}057$). This analysis yielded the most optimal fit indices. Both GFI and CFI met the cut-off criterion of close to .95; this indicated that the model adequately explained the variance and covariance in the ELL student scores. The SRMR of .038 and the RMSEA of .034 met the criteria of less than .08 and .06, respectively; this indicated a good fit between the model-specified variance/covariance matrix and the population variance/covariance matrix and an adequate model specification. Further, all parameter values in the model were statistically nonzero ($p < .001$), thus supporting adequate model specification. Figure 1 summarizes the final structural model estimates in the Combined Sample, including standardized regression weights and latent factor correlations.

Psychometric properties of the SILL–ELL Student Form. Table 2 summarizes the psychometric properties of the SILL–ELL Student Form evaluated in the Combined Sample. The overall Cronbach's alpha was .90, and subscale alphas were: .77 (memory strategies; 7 items); .72 (metacognitive strategies; 4 items); .63 (cognitive strategies; 5 items); .63 (compensation strategies; 5 items); .71 (affective strategies; 3 items); and .75 (social strategies; 4 items). The reliability statistics' range indicated moderate to high level of internal consistency and represented an improvement over studies with similar age populations using both SILL (Gunning, 2011; Magogwe & Oliver, 2007; overall alphas were .82 and .82–.84, respectively) and task-based strategy measures (G. Hu et al., 2009; alpha range: .38–.72).

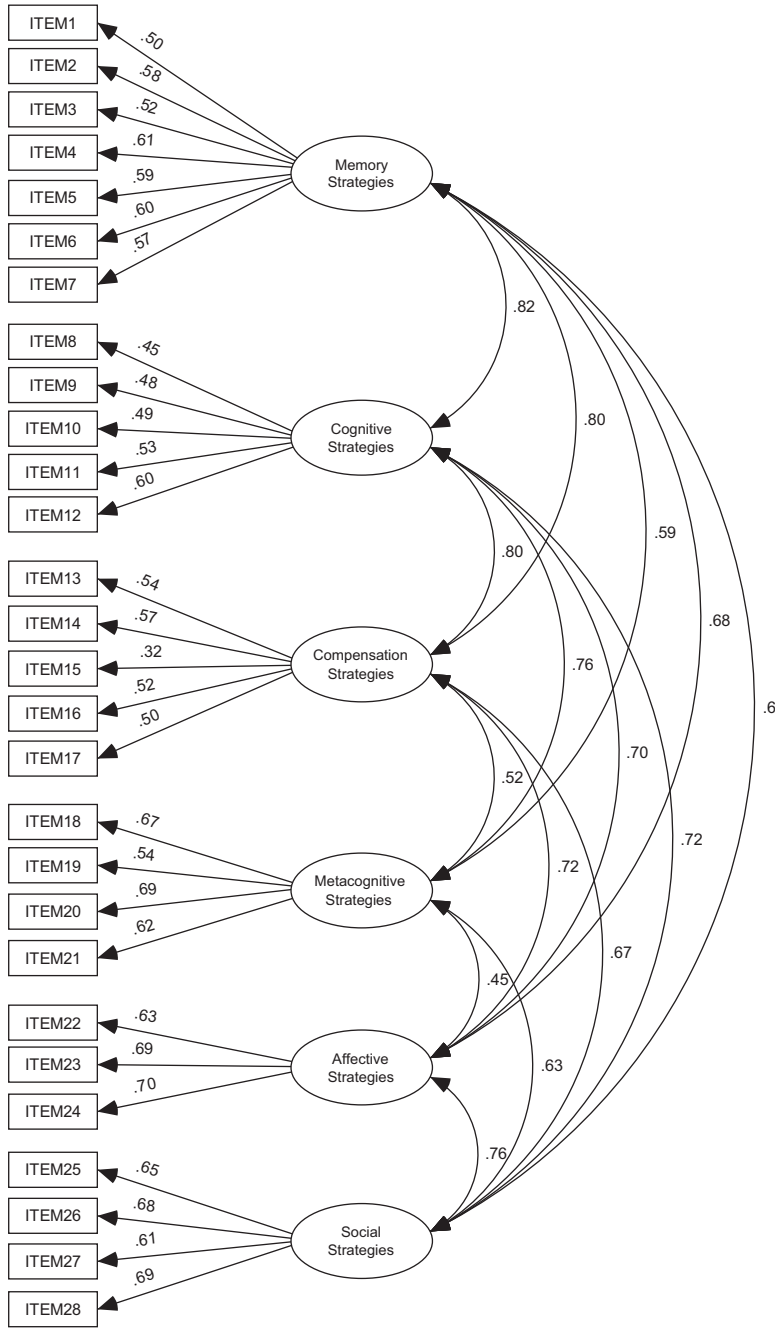
DISCUSSION

Validating the SILL for determining strategy use on the part school-aged ELLs is of great interest given (a) continued interest in examining relationships among holistically measured LLS and student language, academic, and social/behavioral outcomes, (b) the persistent popularity of the SILL in assessing LLS, and (c) the increase of the school-aged ELL population in the United States and other English-speaking countries. A two-stage modification of Oxford's (1990) 50-item adult-oriented ESL/EFL measure of LLS resulted in a shorter, 28-item version of the instrument. The CFA results from validation and combined samples indicate that the new instrument has psychometrically sound characteristics as a measure of LLS among elementary, middle, and high school ELL students. To reflect substantial modifications from the original instrument and to capture the measure's new intended audiences, the modified instrument was entitled SILL–ELL Student Form. The instrument advances the field of LLS research and has practical implications for both researchers and educators.

From a research perspective, three benefits can be identified. First, this study addresses a lack of SILL validation research for school-aged ELLs. Second, the study responds to some theoretical concerns raised in the literature with strategy categorization, namely, strategy category overlap, item appropriateness for second versus foreign language contexts, and consistency of item specificity levels (Cohen, 1998; Hsiao & Oxford, 2002). Third, the study used CFA, a powerful statistical test that addresses potential limitations associated with typically lower reliability estimates when the instrument is administered through a nonnative language to heterogeneous (multilanguage) groups of language learners (Oxford & Burry–Stock, 1995). In their statistical simulation study, Little, Lindenberger, and Nesselroade (1999) found that CFA "accurately corrects for the construct's low measurement quality" and, importantly, "yields unbiased estimates of its relations with other constructs" (p. 207). The latter advantage portends potential benefits of CFA applications to future LLS research, particularly in the light of inconsistencies in previous research (Gardner et al., 1997; Nisbet et al., 2005) and recent evidence suggesting that age and length of L2 study may have a moderating effect on the relationship between L2 outcomes and LLS, in particular, cognitive, memory, compensation, social, and affective strategies (Ardasheva, 2011). This highlights a need to investigate structural relationships among LLS and other individual difference variables and suggests that, because an adequate fit of measurement models is prerequisite for valid interpretations of structural relationships tests (Kline, 2005), the new measure with its psychometrically sound characteristics may contribute to future CFA research.

From a practical point of view, the SILL–ELL Student Form has several advantages for L2 learners and classrooms. First, the instrument can serve as a self-assessment tool for the students as well as a diagnostic or consciousness-raising tool (Oxford, 1990; Vandergrift et al., 2006) by

FIGURE 1
Confirmatory Factor Analysis Results for the SILL–ELL Student Form in Combined Sample ($N = 1{,}057$): Standardized Regression Weights and Latent Factor Correlations



Note. All coefficients are significant at $p < .001$. Individual items are listed in Appendix A.

language and content-area teachers, particularly those implementing instructional models that incorporate explicit strategy instruction (e.g., *Cognitive Academic Language Learning Approach:* Chamot, 2009; *Sheltered Instruction Observation Protocol:* Echevarria, Vogt, & Short, 2007). Such applications, Vandergrift et al. (2006) have noted, could positively influence student L2 attitudes,

TABLE 2
Descriptive Statistics and Psychometric Properties of the SILL–ELL Student Form: Combined Sample

Strategy categories	*M*	*SD*	α	Range		Skewness	Kurtosis
				Potential	Actual		
Memory	3.12	0.85	.77	1–5	1.0–5.0	−0.19	−0.45
Cognitive	3.45	0.84	.63	1–5	1.0–5.0	−0.21	−0.51
Compensation	3.05	0.87	.63	1–5	1.0–5.0	−0.00	−0.62
Metacognitive	4.01	0.82	.72	1–5	1.0–5.0	−0.84	0.15
Affective	2.76	1.10	.71	1–5	1.0–5.0	0.15	−1.04
Social	3.45	1.04	.75	1–5	1.0–5.0	−0.33	−0.73
Overall SILL	3.30	0.69	.90	1–5	1.3–5.0	−0.09	−0.42

Note. $N = 1{,}057$.

self-regulation, and instructional effectiveness. An added plus is that the instrument allows for accurate diagnosis of each strategy category individually, and, thus, can facilitate the development of targeted strategy instruction interventions appropriate for particular instructional contexts and learning goals whose combination determines not only "the types of learning tasks" but also "the types of learning strategies that can be expected to best assist learning" (Chamot, 2009, p. 57). If the development of social language skills is the target, Chamot points out, social and affective strategies would be helpful; if preparation for a grammar or vocabulary test is the goal, memory strategies would be appropriate; the development of academic language skills, particularly in reading and writing, in turn, may be best supported by a repertoire of cognitive and metacognitive strategies. Second, administering the shorter version requires less classroom time and is less demanding on young and early adolescent ELL students. Finally, because the instrument, administered in English to nonnative speakers of English, showed good psychometric properties, its use could be particularly helpful in contexts in which ELLs speak multiple native languages and where the instrument's translation may not be feasible.[11] However, for the instrument to produce quality data and to allow for valid inferences regarding strategy use and needs among school-aged ELLs, its administration should be supported by necessary language accommodations that are appropriate to the student's L2 proficiency level.

CONCLUSION

This study modified and validated Oxford's (1990) SILL, the 50-item adult-oriented ESL/EFL measure of LLS, for use with school-aged ELL students. The results support the validity of the modified, 28-item version of the instrument, which we named SILL–ELL Student Form. This measure showed psychometrically sound characteristics and can enable researchers and educators to diagnose, with an appropriate grain size, LLS use and needs among elementary, middle, and high school ELL students. Each of the strategy categories may have differential impact on student learning depending on the developmental needs and English proficiency level of the individual, the outcome of interest (i.e., linguistic, academic, or cognitive/behavioral), as well as the specific learning and teaching goals and tasks. Thus, this instrument presents an opportunity to advance LLS research and to fine tune and target strategy instruction to individual students and individual classrooms as is most appropriate for their instructional contexts and learning goals.

ACKNOWLEDGMENTS

The authors would like to thank Leo van Lier, Heidi Byrnes, and the anonymous *Modern Language Journal* reviewers for their insightful comments. The authors are also grateful to the participating school district's English-as-a-Second-Language program office for the highly professional assistance with data collection. Any shortcomings of the study are our sole responsibility.

NOTES

[1] In educational psychology, a recent trend is to include the discussion of learning strategies under the umbrella of self-regulatory learning (Dörnyei, 2003, 2005; see also Oxford, 2011).

[2] In his later work, Cohen (1998) proposed that both language learning and language use strategies can be differentiated as cognitive, metacognitive, social, or affective.

[3] These included: (a) null, no relationships model; (b) six correlated models with one to six first-order

factors (general strategy [1], direct–indirect [2], cognitive–metacognitive–socioaffective [3], cognitive–metacognitive–social–affective [4], cognitive–memory–compensation–metacognitive–socioaffective [5], and cognitive–memory–compensation–metacognitive–social–affective [6]); (c) two correlated models extending on model 6 with one and two second-order factors (general strategy [7] and direct–indirect strategy [8], respectively); and (d) six uncorrelated models extending on models 1 through 6.

[4] It is worthwhile to note, however, that the LLS literature has not—as of yet—established the direction of causality (e.g., Ardasheva, 2011; Gardner et al., 1997).

[5] Notably, none of the six EFA analyses reported in Oxford and Burry–Stock (1995) found a six-factor solution: Although common factors across all samples were found, some factors were unique for individual cultural groups including language learners investigated in Puerto Rico, Taiwan, China, Japan, Egypt, and the United States. However, only one of the reviewed studies had a sample size sufficient for EFA analysis.

[6] The two fit indices reported, the Comparative Fit Index (CFI = .75) and the Non-Normed Fit Index (NNFI = .73), were well below their cut-off criteria of close to .95 (Bentler, 1990).

[7] Correlations were particularly high for elementary ($r = .70$) and secondary ($r = .56$) students; this relationship was less pronounced for older and higher-English-proficiency students.

[8] Tseng et al. (2006) have criticized the original 50-item SILL for qualifying the true-of-me scale by frequency adverbs (i.e., never, always) and for using "*behavioral items*" (e.g., using flashcards), thus limiting (a) the instrument's precision in measuring "the *extent* of the correspondence between the item and the learner" and (b) undermining the linear relationship between the item and the latent construct (i.e., "one can be a good memory strategy user in general while scoring low on some of the items," italics in the original; p. 83). This, the authors argued, prevented the possibility of computing cumulative scores on the SILL subscales (see also Dörnyei, 2005). However, the examination of central tendency and variability statistics by individual (retained) SILL items displayed in Appendix A, suggests that the SILL scale provides an adequate range to allow learners selecting a point on a continuum that best corresponds to their behavior. Further, the use of behavioral items to measure latent constructs including learner characteristics and behaviors (e.g., adoption of new learning tools: Lau & Woods, 2009; acculturation: Rubenfeld, Sinclair, & Clément, 2007; and reading amount: Becker, McElvany, & Kortenbruck, 2010) is a common practice in educational research that has produced valid results with both adult and young populations. With younger learners, in particular, behavioral items may be more appropriate from a developmental perspective. Based on these empirical and theoretical considerations, as well as on the CFA analysis results reported later in the paper, we argue that the SILL subscales are appropriate for pooling items into cumulative scores.

[9] Other structural equation modeling applications include alternative model comparisons and model generation (Byrne, 2001; Kline, 2005).

[10] According to Byrne (2001), another potential source of error may arise from respondent characteristics such as social desirability or unwillingness to provide truthful responses.

[11] The district where data for this study were collected, for instance, typically services speakers of over 70 first languages.

REFERENCES

Anderson, J. R. (1982). Acquisition of cognitive skills. *Psychological Review*, *89*, 369–403.

Anderson, J. R. (1989). A theory of human knowledge. *Artificial Intelligence*, *40*, 313–351.

Arbuckle, J. L. (2008). Amos (Version 17.0) [Computer software]. Chicago, IL: SPSS.

Ardasheva, Y. (2011). English language learners in focus: Predictors of English proficiency and academic achievement. *Dissertation Abstracts International: Section A*, *72*, (06).

Ardasheva, Y., & Tretter, T. R. (in press). Contributions of individual differences and contextual variables to reading achievement of English learners: An empirical investigation using Hierarchical Linear Modeling. *TESOL Quarterly*, First published online: October 24, 2012.

Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology*, *102*, 773–785.

Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238–246.

Byrne, B. M. (2001). *Structural equation modeling with Amos: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum.

Byrne, B. M. (2010). *Structural equation modeling using AMOS. Basic concepts, applications, and programming*, (2nd ed.). New York: Routledge.

Chamot, A. U. (2007). Accelerating academic achievement of English language learners: A synthesis of five evaluations of the CALLA Model. *Springer International Handbooks of Education*, *15*, 317–331.

Chamot, A. U. (2009). *The CALLA handbook: Implementing the Cognitive Academic Language Learning Approach* (2nd ed.). New York: Addison–Wesley.

Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach.* Reading, MA: Addison–Wesley.

Chamot, A. U., Dale, M., O'Malley, J. M., & Spanos, G. A. (1992). Learning and problem solving strategies of ESL students. *Bilingual Research Journal*, *16*, 1–38.

Chen, M.–L. (2009). Influence of grade level on perceptual learning style preferences and language learning strategies of Taiwanese English as a foreign language learners. *Learning and Individual Differences, 19*, 304–308.

Clarke, M. A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning, 29*, 121–150.

Clarke, M. A. (1980). The short circuit hypothesis of ESL reading—or when language competence interferes with reading performance. *Modern Language Journal, 64*, 203–209.

Cohen, A. D. (1998). *Strategies in learning and using a second language*. New York: Addison Wesley Longman.

Cohen, A. D., & Macaro, E. (Eds.), (2007). *Language learner strategies: 30 years of research and practice*. Oxford: Oxford University Press.

Cohen, A. D., Weaver, S. J., & Li, T.–Y. (1996). *The impact of strategies-based instruction on speaking a foreign language*, (Working Papers Series No 4). Minneapolis: University of Minnesota, Center for Advanced Research on Language Acquisition.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 2. Literacy*, (2nd ed. pp. 71–83). New York: Springer.

Dillman, D. A. (2007). *Mail and internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*, 3–32.

Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Erlbaum.

Dreyer, C., & Oxford, R. L. (1996). Learning strategies and other predictors of ESL proficiency among Afrikaans speakers in South Africa. In R. Oxford (Ed.), *Language learning strategies around the world* (Tech. Rep. No 13, pp. 62–74). Honolulu: University of Hawaii, Second Language Teaching & Curriculum Center.

Echevarria, J., Vogt, M., & Short, D. J. (2007). *Making content comprehensible for English learners: The SIOP Model* (3rd ed.) Boston, MA: Pearson.

Gardner, R. C., Tremblay, P. F., & Masgoret, A.–M. (1997). Towards a full model of second language learning: An empirical investigation. *Modern Language Journal, 81*, 344–362.

Genesee, F., Lindholm–Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students Placed at Risk, 10*, 363–386.

Goldenberg, C. (2008). Teaching English language learners: What the research does and does not say. *American Educator, 33*, 8–23.

Graham, S., & Macaro, E. (2008). Strategy instruction in listening for lower-intermediate learners of French. *Language Learning, 58*, 747–783.

Griffiths, C. (2007). Language learning strategies: students' and teachers' perceptions. *ELT Journal, 61*, 91–99.

Gunning, P. (1999). *The learning strategies of beginning ESL learners at primary levels*. (Master's thesis, Concordia University, Quebec, Canada, 1997). *Masters Abstracts International*, 37, 1580.

Gunning, P. (2011). *ESL strategy use and instruction at the elementary school level: A mixed methods investigation*. (Unpublished doctoral dissertation). McGill University, Montreal, Quebec, Canada. Retrieved December 10, 2011, from http://digitool.Library.McGill.CA:80/R/-?func=dbin-jump-full&object_id=103480&silo_library=GEN01

Hassan, X., Macaro, E., Mason, D., Nye, G., Smith, P., & Vanderplank, R. (2005). *Strategy training in language learning: A systematic review of available research*. London: University of London, EPPI–Centre.

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

Hsiao, T.–Y. (1997). A study of language learning strategies by college EFL students: A theoretical perspective. *Studies in English Language and Literature, 2*, 29–46.

Hsiao, T.–Y., & Oxford, R. (2002). Comparing theories of language learning strategies: A confirmatory analysis. *Modern Language Journal, 86*, 365–383.

Hu, G., Gu, Y. P., Zhang, J. L., & Bai, R. (2009). *English language learning strategies in Singapore primary school*. (Final Research Report for Project No. CPR 3/03GYQ). Retrieved December 2, 2011, from Centre for Research in Pedagogy and Practice website: http://repository.nie.edu.sg/jspui/bitstream/10497/4173/4/CRP3_03GYQ_Final-ResRpt.pdf

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

Kaufman, D., & Crandall, J. (Eds.). (2005). *Content-based instruction in primary and secondary school settings*. Alexandria, VA: Teachers of English to Speakers of Other Languages.

Kaylani, C. (1996). The influence of gender and motivation on EFL learning strategy use. In R. Oxford (Ed.), *Language learning strategies around the world: Cross-cultural perspectives*, (Tech. Rep. No 13, pp. 75–88). Honolulu: University of Hawaii, Second Language Teaching & Curriculum Center.

Kim, Y.–H., & Jang, E. E. (2009). Differential functioning of reading subskills on the OSSLT for L1 and ELL Students: A multidimensionality model-based DBF/DIF approach. *Language Learning, 59*, 825–865.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*, (2nd ed.). New York: Guilford Press.

Lan, R., & Oxford, R. L. (2003). Language learning strategy profiles of elementary school students in Taiwan. *International Review of Applied Linguistics in Language Teaching*, *41*, 339–379.

Lau, S.–H., & Woods, P. C. (2009). Understanding learner acceptance of learning objects: The roles of learning object characteristics and individual differences. *British Journal of Educational Technology*, *40*, 1059–1075.

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, *4*, 192–211.

Macaro, E. (2006). Strategies for language learning and strategy use: Revising the theoretical framework. *Modern Language Journal*, *90*, 320–337.

MacIntyre, P. D., & Noels, K. A. (1996). Using social-psychological variables to predict the use of language learning strategies. *Foreign Language Annals*, *29*, 373–386.

Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age, and self-efficacy beliefs: A study of language learners in Botswana. *System*, *35*, 338–352.

Marsh, H. W., & Hau, K. T. (1996). Assessing goodness of fit: Is parsimony always desirable?. *Journal of Experimental Education*, *64*, 364–390.

McDonough, S. (1999). Learner strategies. *Language Teaching*, *32*, 1–18.

Montes, F. (2002). Enhancing content area through a Cognitive Academic Language Learning based collaborative in South Texas. *Bilingual Research Journal*, *26*, 697–716.

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

Nunan, D. (1997). Strategy training in the language classroom: An empirical investigation. *RELC Journal*, *28*, 56–76.

O'Malley, M. J., Chamot, A. U., & Walker, C. (1987). Some applications of cognitive theory to second language acquisition. *Studies in Second Language Acquisition*, *9*, 287–306.

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

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

Oxford, R. L. (1999). Relationship between second language strategies and proficiency in the context of learner autonomy and self-regulation. *Revista Canaria de Estudios Ingleses [Canadian Journal of English Studies]*, *38*, 109–126.

Oxford, R. L. (2011). Research timeline: Strategies for learning a second or foreign language. *Language Teaching*, *44*, 167–180.

Oxford, R. L., & Burry–Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL). *System*, *23*, 1–23.

Oxford, R. L., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *Modern Language Journal*, *73*, 291–301.

Oxford, R., Cho, Y., Leung, S., & Kim, H.–J. (2004). Effect of the persistence and difficulty of task on strategy use: An exploratory study. *International Review of Applied Linguistics in Language Teaching*, *42*, 1–47.

Peacock, M., & Ho, B. (2003). Student language learning strategies across eight disciplines. *International Journal of Applied Linguistics*, *13*, 179–200.

Raykov, T., & Marcoulides, G. A. (2006). Fundamentals of structural equation modeling. In T. Raykov & G. A. Marcoulides (Eds.), *A first course in structural equation modeling* (pp. 1–55). Mahwah, NJ: Lawrence Erlbaum.

Rivera–Mills, S. V., & Plonsky, L. (2007). Empowering students with language learning strategies: A critical review of current issues. *Foreign Language Annals*, *40*, 535–548.

Rubenfeld, S., Sinclair, L., & Clément, R. (2007). Second language learning and acculturation: The role of motivation and goal content congruence. *Canadian Journal of Applied Linguistics*, *10*, 309–323.

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*, *11*, 118–131.

Schmidt, R., & Watanabe, Y. (2001). Motivation, strategy use, and pedagogical preferences in foreign language learning. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 313–359). Honolulu, HI: University of Hawaii Press.

Schoonen, R., Hulstijn, J., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in grades 6, 8, and 10. *Language Learning*, *48*, 71–106.

Shevlin, M., & Miles, J. N. V. (1998). Effects of sample size, model specification, and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences*, *25*, 85–90.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.) Mahwah, NJ: Lawrence Erlbaum.

Takeuchi, O. (1993). Language learning strategies and their relationship to achievement in English as a foreign language. *Language Laboratory*, *30*, 17–34.

Tseng, W.–T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*, 78–102.

Vandergrift, L., Goh, C., Mareschal, C., & Tafaghodtari, M. H. (2006). The Metacognitive Awareness Listening Questionnaire (MALQ): Development and validation. *Language Learning, 56*, 431–462.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning, 60*, 470–497.

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

APPENDIX A

Strategy Inventory for Language Learning (SILL)–ELL Student Form: Descriptive Statistics by Individual Items

Strategies	*M*	*SD*	Mode
Memory			
1. I use flashcards to learn new English words.	2.60	1.35	1
2. I use rhymes to help me learn new English words.	2.82	1.37	3
3. I act out new English words.	2.65	1.42	1
4. I use new English words in a sentence to help me learn them.	3.65	1.23	5
5. I learn new words by thinking about when I can use them.	3.29	1.25	3
6. When I hear a new English word I think of a picture to help me learn the word.	3.45	1.32	5
7. I learn new words by thinking about where I first saw them on the page, on the board, or on a street sign.	3.35	1.27	3
Cognitive			
8. I read for fun in English.	3.51	1.35	5
9. I first read a page (a text) quickly and then go back and read it carefully.	3.40	1.28	3
10. I look for words in English that are like my own language.	3.39	1.43	5
11. I break long words into small parts to figure out what they mean.	3.55	1.24	5
12. I make summaries of things I hear or read in English.	3.38	1.27	3
Compensation			
13. If I can't think of an English word, I show what I mean with my hands.	2.90	1.41	3
14. I make up a new word if I can't think of an English word.	2.80	1.45	1
15. When I read in English, I don't look up every new word in a dictionary.	2.95	1.38	3
16. I try to guess (predict) what people will say next in English.	3.03	1.36	3
17. If I can't think of an English word, I use a word that means the same thing.	3.56	1.23	5
Metacognitive			
18. I see my English mistakes and try to do better.	4.03	1.10	5
19. I listen well (carefully) when people speak English.	4.09	1.08	5
20. I look for ways to be a better student of English.	4.23	1.07	5
21. I think about how well I am doing in English.	3.68	1.20	5
Affective			
22. I give myself a gift or a treat when I do well in English.	2.78	1.49	1
23. I write about how I feel when I am learning English in my journal.	2.67	1.54	1
24. I talk to people about how I feel when I am learning English.	2.82	1.47	1
Social			
25. If I don't understand, I ask English speakers to slow down or say it again.	3.73	1.27	5
26. I ask English speakers to correct me when I talk.	3.18	1.47	5
27. I practice English with other students.	3.42	1.44	5
28. I ask for help from English speakers.	3.46	1.33	5

Note. The actual range on all individual items was 1–5. Only SILL–ELL Student Form items (those retained from the original 50-item SILL) are included in Appendix A.

APPENDIX B

Items Deleted from the Original SILL

Memory strategies

1. I go over English lessons often.
2. I use what I already know to help me learn new things in English.

Cognitive strategies

3. I say or write new English words many times to help me learn them.
4. I try to speak English like Americans.
5. I work on my English sounds to make them better.
6. I use the English words I know in different ways.
7. I often start talking in English with others.
8. I watch TV shows and movies in English.
9. I write notes, emails, letters, or reports in English.
10. I try to find patterns in English. (For example: he reads/she reads.)
11. I try to figure out what the whole sentence means, not word by word.

Compensation strategies

12. I try to guess what new English words mean.

Metacognitive strategies

13. I find many ways to use my English.
14. I plan my day so I will have time to study English.
15. I look for people I can talk to in English.
16. I look for things to read in English.
17. I have clear goals for making my English better.

Affective strategies

18 I try to calm down (relax) when I feel scared of using English.

19 I don't let my English mistakes stop me from speaking.

20 I can tell if I am scared of using English.

Social strategies

21. I ask questions in English.
22. I try to learn about Americans and how they live.