Applied Linguistics 27/1: 78–102
doi:10.1093/applin/ami046

A New Approach to Assessing Strategic Learning: The Case of Self-Regulation in Vocabulary Acquisition

WEN-TA TSENG, ZOLTÁN DÖRNYEI, and NORBERT SCHMITT

University of Nottingham

This article draws on work done in educational psychology to propose a new approach to generating a psychometrically-based measure of second language learners' strategic learning, operationalized as their *self-regulatory capacity*, as an alternative to the scales traditionally used to quantify language learning strategy use. The self-regulation instrument was developed through a three-phase process, focusing on the realm of vocabulary learning. The first phase involved the generation of an item pool, the second a pilot study in a sizeable sample, and the third an evaluation of the psychometric properties of the revised instrument, using confirmatory and exploratory factor analysis. The results show that the proposed instrument has satisfactory psychometric characteristics and that the hypothesized theoretical model had a good fit with the data. We argue that the results provide evidence for the validity of transferring the theoretical construct of self-regulation from educational psychology to the area of second language acquisition. We also propose that instruments targeting learner self-regulation in a similar way to the questionnaire presented in this study can provide a more psychometrically sound measure of strategic learning than traditional language learning strategy scales.

INTRODUCTION

The past twenty years have witnessed a large body of second language research targeting *language learning strategies* (e.g. Anderson 2003; Chamot *et al*. 1999; Cohen 1998, 2002; Ehrman *et al*. 2003; Grenfell and Harris 1999; Lan and Oxford 2003; Macaro 2001; MacIntyre 1994; McDonough 1995, 1999; Nunan 1997; O'Malley *et al*. 1985; O'Malley and Chamot 1990; Oxford 1990, 1996; Purdie and Oliver 1999; Purpura 1999; Wenden 1991; Wenden and Rubin 1987; Yamamori *et al*. 2003). While some of this research has explicitly sought to push the theoretical understanding of language learning strategies forward, the majority of the work in the learning strategy literature had more practical goals, namely to explore ways of empowering language learners to become more self-directed and effective in their learning. In general, strategy specialists believe that learners with strategic knowledge of language learning, compared with those without, become more efficient, resourceful, and flexible, thus acquiring a language more easily. The suggestion is that if learners can develop, personalize, and use a

repertoire of learning strategies, they will be able to achieve language proficiency in a much facilitated manner. Indeed, most researchers would agree with Macaro's (2001: 264) conclusion that 'One thing seems to be increasingly clear and that is that, across learning contexts, those learners who are pro-active in their pursuit of language learning appear to learn best.'

While we consider this appealing notion about the importance of strategic learners to be true, research into language learning strategies has unfortunately suffered from a series of problems, which weaken the theoretical basis for such a conclusion. These problems stem partly from definitional fuzziness, and partly from the psychometric properties of the assessment instruments used (e.g. Ellis 1994; Skehan 1989; for recent reviews, see Dörnyei 2005; Dörnyei and Skehan 2003). Indeed, as a result of similar concerns, in educational psychology the research boom in the area of learning strategies characterizing the1980s had almost completely petered out by the mid-1990s, and the term 'learning strategy' now rarely appears in research publications. Nevertheless, the notion of strategic learning is still of great interest for theoreticians and practitioners alike, and this paper is intended to contribute to the ongoing discussion by trying to achieve two main goals.

First, we will outline a new approach to conceptualizing and assessing strategic learning, based on the notion of *self-regulation* drawn from the field of educational psychology. As we will argue, rather than focusing on the outcomes of strategic learning (i.e. the actual strategies and techniques the learners apply to enhance their own learning), this conceptual approach highlights the importance of the learners' innate self-regulatory capacity that fuels their efforts to search for and then apply personalized strategic learning mechanisms. That is, in line with contemporary theories of self-regulation in educational psychology, our approach targets the core learner difference that distinguishes self-regulated learners from their peers who do not engage in strategic learning.

Second, we will put the proposed theoretical approach to the test by devising an instrument that operationalizes the newly-conceived system of self-regulatory capacity. We offer detailed guidelines on the development and the validation of this new strategic learning measure and the Appendix contains the whole questionnaire. An important aspect of our approach is that it is situated in one specific language learning domain, vocabulary learning. The reason for this focus has partly been the significance that we attach to the mastery of lexis in the second language acquisition process, and partly our general concern that the term 'language learning' is too broad and is often used to encompass diverse acquisitional processes depending on the particular targeted language area in question. Thus, for the sake of conceptual clarity, we have decided to model the self-regulatory system proposed in this paper on one particular language learning domain, vocabulary learning, but at the same time we offer a detailed description

of the procedures used to develop our instrument so that this can serve as a template for other content areas as well. Thus, we believe that our suggested approach is transferable to researching other facets of second language learning.

BACKGROUND

Language learning strategies and self-regulation

A major underlying problem with strategy research relates to the diverse conceptualizations of 'learning strategies'. In the past, there was no coherent agreement on exactly what the defining criteria for language learning strategies are and regrettably the situation remains the same today. A noteworthy debate, for instance, is whether learning strategies should be regarded as either observable behaviours or inner mental operations, or both (cf. Dörnyei 2005; Ellis 1994), an issue that is not restricted to L2 research but also concerns the broader field of educational psychology. To illustrate this, let us take a standard definition from the 1980s: Weinstein and Mayer (1986: 315) define strategies as 'the behaviours and thoughts that a learner engages in during learning that are intended to influence the learner's encoding process'. Recently, Weinstein *et al.* (2000: 727) have redefined learning strategies as 'any thoughts, behaviours, beliefs, or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills'. It is clear that from a scientific point of view a phenomenon is highly unlikely to be both behavioural and cognitive in nature, and yet it seems that rather than sorting out the cognition/behaviour issue, the scope of learning strategies has been further broadened. It needs little justification that a concept cannot be conceptualized as thought, belief, emotion, and observable behaviour at the same time; however, this conceptual ambiguity is less of a problem in educational psychology, because, as we will see below, the term 'learning strategy' has virtually been abandoned for research purposes and has been maintained primarily for pedagogical discourse only.

A full discussion of the strategy controversy goes beyond the current scope of this paper (for a comprehensive review, see Dörnyei 2005), yet to summarize the main question in a nutshell, the real problem facing researchers is how to distinguish strategic learning from 'ordinary' learning. Weinstein *et al.* (2000) offers three critical characteristics of strategic learning: it is *goal-directed, intentionally invoked*, and *effortful*. The problem with these intuitively appealing attributes is that they can also be true about *motivated* or *hard* learning in general, without any 'strategic' element. Cohen (1998) highlights a further important aspect of learning strategies, the element of *choice*: he argues that it is an essential feature of these strategies that they are voluntarily employed by the learner. While this is clearly

important in distinguishing learning strategies from creative teacher-owned tasks that the learner engages in, choice in itself is still not enough to distinguish strategies from non-strategic learner behaviours because students tend to make several choices concerning their learning process that are obviously not strategic in nature.

Addressing the same issue, Riding and Rayner (1998) have argued that an activity becomes strategic when it is particularly *appropriate* for the individual learner, in contrast to general learning activities which a student may find less helpful. Accordingly, learners engage in strategic learning if they exert purposeful effort to select, and then pursue, learning procedures that they believe will increase their individual learning effectiveness. This, however, means that learning strategies conceptualized in this vein can only be defined relative to a particular agent, because a specific learning activity may be strategic for one and non-strategic for another. We agree with this conclusion and its implication that it is not what learners do that makes them strategic learners but rather the fact that they put creative effort into trying to improve their own learning. This is an important shift from focusing on the product—the actual techniques employed—to the self-regulatory process itself and the specific learner capacity underlying it. By the beginning of the 1990s, educational psychologists had gone through this transformation process and the study of self-regulation had come of age, causing a 'virtual explosion of work in this area' (Zeidner *et al*. 2000: 750), thereby becoming a 'natural and organic part of the landscape of psychology and education' (Zeidner *et al*. 2000: 749). We feel it is time to draw upon this explosion of work and to apply it, if possible, in a second language context.

To conclude this brief overview, although the conceptual shift outlined above has not solved the problem of what learning strategies (or 'self-regulatory mechanisms' as they came to be called) actually are, it has resulted in the broadening of the perspective, with self-regulation proposed to be made up of a whole series of integrated and interrelated micro-processes, of which learning strategy use is only one. Other components, based on Kuhl and Goschke (1994), Winne and Perry (2000), and Zeidner *et al*. (2000), include goal setting, strategic planning, action plans and action schemata, monitoring and metacognition, action control and volitional control mechanisms, strategic tactics and operations, effective time management, self-motivational beliefs (self-efficacy, outcome expectations, intrinsic interest, and goal orientation, etc.), evaluation and self-reflection, receiving and processing feedback, experiencing pride and satisfaction with one's efforts, and establishing a congenial environment.

Thus, the broader concept of self-regulation has created considerably more 'leeway' for researchers and even if 'self-regulatory mechanisms' carry the same problems as 'learning strategies', this insufficient understanding does not prevent researchers from making headway in understanding other aspects of self-regulation.

Assessing strategic learning

If we borrow the theoretical construct of self-regulation from educational psychology, it still leaves the problem of how to operationalize and measure it. Learning strategy use and, more generally, strategic learning, have typically been measured by self-report questionnaires in the past, since strategic learning is driven by mental processes that do not often lend themselves to direct observation and, therefore, for an accurate assessment of the extent of their functioning we need to draw on the learners' own accounts. We must note at this point that increased utilization of stimulated recall methodology (cf. Gass and Mackey 2000) offers a promising future research direction in this area, but so far little research has been done in this vein.

The existing self-report questionnaires in this area are based on the assumption that strategy use and strategic learning are related to an underlying *trait* because items ask respondents to generalize their actions across situations rather than referencing singular and specific learning events (Winne and Perry 2000). The most famous such instrument in educational psychology is the 'Motivated Strategies for Learning Questionnaire' (MSLQ) questionnaire, developed at the University of Michigan by Paul Pintrich and his colleagues (Pintrich *et al*. 1991; reprinted in VanderStoep and Pintrich 2003). The MSLQ is aimed at college students and, as the name of the instrument indicates, the items cover two broad areas, *motivation* and *learning strategies*—here only the latter part will be discussed. The Learning Strategies category includes 50 items, each using a 7-point Likert scale anchored by 'not at all true of me' (1) and 'very true of me' (7), and is divided into two sections: (a) 'Cognitive and Metacognitive Strategies', which includes subscales labelled *rehearsal*, *elaboration*, *organisation*, *critical thinking*, and *metacognitive self-regulation*; and (b) 'Resource Management Strategies', which includes the subscales of *time and study environment*, *effort regulation*, *peer learning*, and *help seeking*. All these subscales are cumulative in the sense that composite subscale scores are formed by computing the means of the individual item scores in a subscale.

In the L2 field, the most frequently used instrument for assessing language learning strategy use during the past decade has been the 'Strategy Inventory for Language Learning' (SILL). It has been developed by Rebecca Oxford (1990) based on her strategy taxonomy and thus it consists of six scales: (a) 'Remembering more effectively' (memory strategies); (b) 'Using your mental processes' (cognitive strategies); (c) 'Compensating for missing knowledge' (compensation strategies); (d) 'Organising and evaluating your learning' (metacognitive strategies); (e) 'Managing your emotions' (Affective strategies); and (f) 'Learning with others' (social strategies). Scale scores are obtained, similarly to the MSLQ, by computing the average of the item scores within a scale. The author has published two versions of the instrument,

one for speakers of English learning other target languages (80 items) and one for learners of English as an L2 (50 items).

The SILL items all involve five-point rating scales ranging from 'never or almost never true of me' to 'always or almost always true of me'. At first sight, these scales are similar to the scales used in the MSLQ discussed above, but a closer look reveals two fundamental differences. First, although both scale types use the term 'true of me', the MLSQ scales range from '*not at all*' to '*very*' whereas the SILL scales from '*never or almost never*' to '*always or almost always*'. Second, the items themselves are of a different nature. The items in the MSLQ are general declarations or conditional relations focusing on general and prominent facets of the learning process (i.e. *when doing this . . . I try to . . .*). On the other hand, the SILL items are more specific, each one more or less corresponding to a language learning strategy. These two changes, however, result in a major difference in the psychometric character of the two inventories. The items in the MSLQ scale tap into general trends and inclinations and can therefore be assumed to be in a linear relationship with some corresponding underlying learner trait. This is further enhanced by the rating scales asking about the *extent* of the correspondence between the item and the learner, answered by marking a point on a continuum between 'not at all' and 'very'. Thus, every attempt has been made to make the items *cumulative*, which is why the scale scores can be formulated by pooling all the scale items. The SILL, on the other hand, focuses on specific strategic behaviours and the scale descriptors indicate *frequencies* of strategy use (ranging between 'never' to 'always'). These items are, therefore, *behavioural items*, which means that we cannot assume a linear relationship between the individual item scores and the total scale scores; for example, one can be a good memory strategy user in general while scoring low on some of the items in the memory scale (e.g. acting out a new word or using flashcards).

Thus, the scales in the SILL are *not* cumulative and computing mean scale scores is *not* justifiable psychometrically. To illustrate the problem in broad terms, a high score on the SILL is achieved by a learner using as many different strategies as possible. Therefore, it is largely the quantity that matters. This is in contradiction with learning strategy theory, which has indicated clearly that in strategy use it is not the quantity but the *quality* of the strategies that is important (cf. the discussion about 'appropriateness' as a critical feature of learning strategies): As an extreme, one can go a long way by using only one strategy that perfectly suits the learner's personality and learning style; and even if someone uses several strategies, it does not necessarily mean that the person is an able strategy user because, as Ehrman *et al*. (2003: 315) have found, 'less able learners often use strategies in a random, unconnected, and uncontrolled manner.' Such qualitative aspects, however, are not addressed by the SILL. It is interesting to read that Oxford

and her colleagues' recent reappraisal of this issue is in accordance with the above argument:

> Low reported strategy use is not always a sign of ineffective learning. Also, reportedly high-frequency use of strategies does not guarantee that the learning is successful. In a casual class observation, one might see some learners working eagerly and using many strategies, but . . . do not employ those strategies effectively. Studies relying solely on frequency data may miss this point. Because frequency results alone do not explain everything about strategy use, it is necessary to include other indices of learners' behaviours that reflect their decision making. 'The more, the better' is not always the case in strategy use. (Yamamori *et al*. 2003: 384)

The same problems also hold in the more specific area of vocabulary learning strategies (VLS). Let us examine three VLS instruments that are relevant to our discussion of the MSLQ and SILL mentioned above. First, Schmitt (1997) based his VLS taxonomy on Oxford's taxonomy, but further refined it for vocabulary learning by dividing the strategies into Discovery Strategies and Consolidation Strategies categories. Schmitt's prime intention has been to bridge the gap between vocabulary learning and language learning strategies and, as a result, his items are written in a behavioural fashion, which he acknowledges leads to quite a large degree of overlap between the categories. The extent to which the two categories of strategies are different from each other is therefore difficult to interpret.

Secondly, Gu and Johnson (1996) compiled a section of vocabulary learning strategies in their Vocabulary Learning Questionnaire (VLQ, Version 3). There are 91 items in total in the section. The main purpose of their research has been to use VLS as the predictor to see its effects on both vocabulary knowledge and general language proficiency. Although Gu and Johnson do not adopt frequency-based evaluations of the items but the ones similar to MSLQ's, it is nevertheless clear that the items concern very specific vocabulary learning behaviours (e.g. *I write down both the Chinese equivalent and the English synonyms of the word I look up*). Like the problems underlying SILL, it seems therefore psychometrically misleading to sum this part of scores as indicative of one's overall capacity to use VLS.

Thirdly, Stoffer's (1995) work—the Vocabulary Learning Strategies Inventory (VOLSI)—attempts not only to classify VLS but also to establish a psychometrically sound instrument. The behavioural nature of the items she used and the evaluative criteria of the instrument were not dissimilar to those adopted in SILL. However, instead of classifying the strategies beforehand, she used factor analysis to determine the categories. Using actual learner data to establish the categories might be considered a step forward, but the danger of this approach is to obtain factors that are specific to the idiosyncratic sample. In the case of the VOLSI this was indeed the case: Many

items for a particular factor seem to be conceptually unrelated. For instance, both '*Break word into its parts (prefix, root)*' and '*Record words on tape and listen*' are grouped into the factor 'Strategies used for Word Analysis'. Clearly, the second strategy has more to do with the decision to use a physical learning aid (tape recorder) than any word analysis approach. We suspect that the sometimes incoherent factor categories stem at least partially from the problem discussed above: the items are based on the frequency of behaviour. It may well be that a factor analysis that is based on an item format more in line with the MSLQ would produce a more satisfactory classification. In sum, the approaches taken by the three studies above can take us only part of the way towards understanding the behaviour students use in learning vocabulary. It seems that the area of VLS is still in need of an instrument which is truly psychometrically valid.

Principles and objectives of the study

Because of the emerging consensus that the approach taken by the SILL needs to be modified to obtain a more accurate measure of strategic learning as well as the fact that the problems underlying the SILL have also been reflected in the area of VLS, we have pursued a research project to conceptualize, develop and test a new instrument. In the light of the theoretical and measurement arguments above, we set the following objectives:

(a) The new instrument should target the learner trait of self-regulatory capacity rather than survey specific behavioural habits, as has been the norm in second language research. Thus, the self-report items should be similar to the MSLQ items in that they constitute general declarations or conditional relations rather than descriptions of specific strategic behaviours.

(b) The structure and content of the new instrument should be based on a theoretical construct. Because of the theoretical problems surrounding the notion of learning strategies and consequently the existing learning strategy taxonomies, we decided to draw on a system of self-regulatory strategies from the area of educational psychology developed by Dörnyei (2001), based on Kuhl's (1987) and Corno and Kanfer's (1993) taxonomies of action control strategies. This system consists of five facets:

- *Commitment control*, which helps to preserve or increase the learners' original goal commitment (e.g. keeping in mind favourable expectations or positive incentives and rewards; focusing on what would happen if the original intention failed).
- *Metacognitive control*, which involves the monitoring and controlling of concentration, and the curtailing of any unnecessary procrastination (e.g. identifying recurring distractions and developing defensive routines; focusing on the first steps to take when getting down to an activity).

- *Satiation control*, which helps to eliminate boredom and to add extra attraction or interest to the task (e.g. adding a twist to the task; using one's fantasy to liven up the task).
- *Emotion control*, which concerns the management of disruptive emotional states or moods, and the generation of emotions that will be conducive to implementing one's intentions (e.g. self-encouragement; using relaxation and meditation techniques).
- *Environmental control*, which helps to eliminate negative environmental influences and to exploit positive environmental influences by making the environment an ally in the pursuit of a difficult goal (e.g. eliminating distractions; asking friends to help and not to allow one to do something).

(c) To further increase the validity of the construct, we have decided to situate the construct in one particular learning domain only, that is vocabulary learning. Acquiring sufficient lexis is a key aspect for developing language skills. Knowing a sufficient number of words is necessary for both daily oral communication and various types of reading (Nation 1990, 2001; Nation and Waring 1997). For instance, around 3,000 spoken word families are required to cover about 96 per cent of the words used in daily conversation (Adolphs and Schmitt 2003). Similarly, 3,000 written word families are required for reading authentic texts (Laufer 1997). Furthermore, to match the lexicon of a native university graduate, a vocabulary size approaching 20,000 word families is needed (Goulden *et al*. 1990). Carter (1998) remarks that non-native speakers need to learn 1,000 word families per year to catch up with the level of an educated native speaker. Although it is not realistic for most foreign learners to fully reach the level of an educated native speaker, it is nevertheless essential for foreign learners to commit themselves to sustained vocabulary study in order to reach the vocabulary requirements for even daily conversation and the modest reading of authentic materials. Given English's very large lexicon, acquiring a vocabulary large enough to cope is probably the major hurdle facing EFL learners (Nation and Meara 2002), and clearly good self-regulation would be an important asset in this major task. Our intention is therefore to develop a self-report instrument that explicitly targets this self-regulatory capacity in the area of English vocabulary learning, assuming that if this attempt is successful, it can, in turn, serve as a model for other areas of language learning as well.

OVERVIEW OF THE INSTRUMENT DEVELOPMENT

With the above-mentioned objectives in mind, we initiated an extended test design process (for a theoretical overview of such a process, see Dörnyei 2003)

that included developing an item pool for each subscale; preparing a first version of the instrument; piloting this version; based on the pilot results designing the final version; and finally administering the instrument to a sample of language learners to validate it. In the rest of the paper we describe each of these steps in detail.

First phase of the study: developing the item pool

From Dörnyei's (2001) work, we had a theoretical basis for believing that self-regulation consists of five facets; therefore the first concrete step in developing our instrument was to write scale items targeting these facets. It is a standard recommendation in questionnaire design theory that the quality of the item pool can be improved by involving the learners themselves in the item-generating process, tapping into their ideas and perceptions (cf. Dörnyei 2003). In order to do this, the first phase of our research included conducting three focus group interviews to inspire the writing of potential items. Following the established practice (Krueger and Casey 2000; Stewart and Shamdasani 1990), each focus group consisted of eight participants (who were all university students in Taiwan), and the composition of these groups aimed at achieving variation in terms of both the gender and the educational background of the participants. The focus groups were conducted in the participants' L1 (i.e. Mandarin), using the questions and probes listed in Appendix 1.

Analysis of the focus group data yielded a total of 36 ideas, which were then translated into instrument items. In addition, we added another nine items based on points derived from our review of the literature. This resulted in a total of 45 items on five subscales:

1 Commitment Control: 12 items
2 Metacognitive Control: 8 items
3 Satiation Control: 8 items
4 Emotion Control: 9 items
5 Environment Control: 8 items

All the questionnaire items involved six-point Likert scales ranging from 'strongly disagree' to 'strongly agree', and respondents were required to mark their answers by ticking the appropriate box for the option that best expressed their personal vocabulary learning experience.

Second phase of the study: piloting the instrument

Before putting the pilot version into effect, the lead author prepared both Chinese and English versions of the instrument. We then invited three translation specialists, all holding Ph.D.s in Speech Communication or Linguistics,[1] to examine the coherence of meanings of the items presented

in both versions. With the help of these consultants, we finalized the instrument and submitted it to extensive piloting to identify items that did not discriminate among the participants successfully or which reduced the internal consistency reliability of the scales.

Participants

The participants of the pilot study were 192 (89 males; 113 females) Taiwanese university students from two universities in Taiwan, including freshmen, sophomores, and juniors. A proportion, 82 participants, were from a private university, and the other 110 participants were from a national university. They were all studying English as a foreign language. Although no proficiency measures were available for the participants, previous research has shown that an average national university freshman student has an English vocabulary size of around 4,000 words (Tseng 2000). However, it is generally held in Taiwan that, on average, students from national universities have better English proficiency than students from private universities, so the students from the private university may have had a vocabulary size somewhat smaller than this. The participants came from a variety of subject backgrounds such as English, business, and education.

Procedures

The lead author contacted the English teachers at several universities in Taiwan to ask for permission to administer the instrument. Two universities, a national and a private one, allowed access to the pilot study, which was conducted in October, 2003. The lead author was present at each research site to explain the purpose of the study and also to make it clear that the result would not have any effect on the students' grades in the courses they were taking. The participants received the Chinese version of the questionnaire and were told that there was no time limit in completing it; they were also encouraged to report the items which were worded inappropriately.

Item analysis

Two kinds of item analysis were conducted: Extreme Group Method and Corrected Item-Total Correlation (Kaplan and Saccuzzo 1993; Dörnyei 2003). With regard to the Extreme Group Method, an item was considered acceptable if it could discriminate well between the total test scores of the upper 33 per cent and the lower 33 per cent of the participants. An independent samples t-test was performed for the purpose of this analysis. With regard to the Corrected Item-Total Correlation Method, an item was considered suspect if the correlation between it and its subscale was below 0.40, which would show that the relationship between item and subscale is

weak (Hatch and Lazaraton 1991). The results of the item analyses showed that four items did not perform well in at least one of the analyses and thus they were deleted, leaving 41 items for the subsequent reliability analysis.

Reliability

The next step was to conduct an internal consistency reliability analysis to determine the reliability of each subscale. We computed Cronbach Alpha coefficients for each subscale and then determined which items would form the most coherent scales. As a result, we retained four items per subscale, making a total of twenty items for the overall self-regulation scale. The final version of the 'Self-Regulating Capacity in Vocabulary Learning' scale (SRCvoc) is presented, in Appendix 2. The reliability results of the final subscales are shown in Table 1. Given that the mean Cronbach Alpha coefficient is 0.78 and all the individual scale coefficients are above 0.70, we can safely conclude that the scales performed well in terms of reliability.

Table 1: The internal consistency reliability of the subscales in the first study

Self-regulation capacity	Cronbach alpha
Commitment control	0.81
Metacognitive control	0.71
Satiation control	0.80
Emotion control	0.82
Environment control	0.74

The third phase of the study: evaulating the final instrument

In the pilot, we successfully screened out the items which did not work well in the item analyses and then reduced the instrument to a total of 20 items, retaining the best four items for each of the five subscales. In the third phase of our project, we administered the revised version of the instrument to a different target sample—senior high school students—to test whether the reliability of the instrument could be established with a different target group and to check the construct validity of our measure.

Participants

The participants of the second study were 172 (90 males; 82 females) Taiwanese senior high school final year students. They came from two public schools (89 and 83 participants from each), which were average schools,

neither particularly strong nor weak. Again, no proficiency measures were available for the participants, but the students of the two schools were considered to have similar English competence at around the pre-intermediate level. Each participant was paid $10 NT dollars when completing the instrument.

Procedures

We applied the procedures of the pilot study to this follow-up study. After gaining access to the research sites, the survey took place in November 2003. The lead author was present in both research sites to clarify the purpose of the study and participants were assured that the result was for research purposes only. No time constraint was imposed on the completion of the questionnaires.

Reliability

Table 2 presents the Cronbach Alpha internal consistency reliability coefficients of the five subscales. It was found that the reliability indices were only marginally lower than in the pilot sample, with a mean scale coefficient of 0.77, which allowed us to conclude that the instrument was well within the range of acceptability as suggested by DeVellis (1991) (see Table 2). This confirmed that the final form of the 'Self-Regulating Capacity in Vocabulary Learning' scale (SRCvoc) is a reliable research instrument.

Table 2: The internal consistency reliability of subscales in the second study

Self-regulation capacity	Cronbach alpha
Commitment control	0.85
Metacognitive control	0.79
Satiation control	0.75
Emotion control	0.78
Environment control	0.66

Confirmatory factor analysis

We explored the construct validity of the instrument by means of confirmatory factor analysis (CFA). We tested the hypothesis that the underlying latent construct of Self-Regulatory Capacity in Vocabulary Learning was a general factor with five subdimensions, as the initial theory would suggest. In the CFA, we adopted a 'confirmatory modelling strategy' (Hair *et al*. 1998), that is, assessed the hypothesized model for its fit to the observed data, using the software AMOS 4.0[2] (Arbuckle and Wothke 1999).

To offer a wide-ranging view of the fit structure of our model, we followed the principles provided by Bagozzi and Yi (1988), approaching the issue from three aspects:

1 Preliminary Fit Criteria[3]
2 Overall Model Fit[4]
3 Fit of Internal Structure of Model.

Regarding the Preliminary Fit Criteria, we checked that the factor loadings were within the acceptable range, that is, between 0.50 and 0.95; the correlations among the variables were not too close to 1.00 (<0.95); and that no large standard errors existed. The results in Table 3 indicate that the five subscales discriminated well with one another, producing no empirical redundancy, and the appropriateness of the factor loadings suggests good acceptability of the construct validity of SRCvoc.

For the Overall Model Fit, we computed a number of standard goodness-of-fit indices, each testing slightly different aspects of the model–data fit: chi-square value, adjusted goodness-of-fit index, incremental fit index, critical N, and the ratio of sample size to number of free parameters. In addition, Hair *et al*. (1998) offered some further indices for measuring overall model fit (e.g. normed chi-square value, Tucker–Lewis Index, and several parsimony indices for alternative models), which we also used, as they served as extra support for the framework provided by Bagozzi and Yi (1988). Table 3 presents a list of these indices with the appropriateness thresholds and the actual results.

Although the chi-square test and χ^2/df index indicated less-than-adequate fit, it has been repeatedly suggested in the literature that the chi-square test may not be accurate with sample sizes over 200 (e.g. Bagozzi and Yi 1988; Schumacker and Lomax 1996); since our sample size was close to 200, we turned to the other fit indices to test the adequacy of our model. Four incremental fit indices (i.e. IFI, NFI, CFI, and TLI) unanimously indicated that the hypothesized model improved very significantly in fit over the null model (i.e. well beyond the recommended value). This information suggested that the hypothesized model was meaningful and appropriate. Extra evidence for this claim was provided by the fact that the GFI was 0.94 and the ratio of sample size to number of free parameters was 17:1. In particular, the GFI suggests that our conceptualized model fits well with the empirical data, as the variance in the sample covariance was well accounted for by the estimated population covariance matrix. Besides, Hoelter's critical *N* (CN) showed the adequacy of the sample size employed in the model in the sense that our sample size was both larger than 68 at 0.05 and 73 at 0.01. Apart from the RMSEA value, the other fit indices also showed that the hypothesized model formed a good overall model fit. On balance, it seemed from the sum of these results that the model did explain the data well.

Acquiring overall model fit is a necessary but insufficient proof of model adequacy, so we also explored whether the model met the criteria of various

Table 3: Summary of the evaluation of measurement model fit

	Levels of acceptable fit	Evaluation of the instrument
1 Preliminary Fit Criteria		
Correlations among variables	Not close to 1.00	Good (0.56∼0.80)
Factor loadings	$0.50 < \lambda < 0.95$	Very good (COM = 0.88; META = 0.85; SAT = 0.84; EMO = 0.88; ENV = 0.69)
Standard errors	No very large standard errors	Good (0.13∼0.15)
2 Overall Model Fit		
Chi-square value	Nonsignificant with p-value > 0.05	Poor ($\chi^2 = 27.89$, $p < 0.001$)
χ^2/df	< 3	Poor (χ^2/df = 5.58)
Goodness of fit index (GFI)	> 0.90	Very good (GFI = 0.94)
Adjusted goodness of fit index (AGFI)	> 0.90	Marginal (AGFI = 0.81)
Incremental fit index (IFI)	> 0.90	Very good (IFI = 0.96)
Normed fit index (NFI)	> 0.90	Very good (NFI = 0.96)
Comparative fit index (CFI)	> 0.90	Very good (CFI = 0.96)
Tucker–Lewis index (TLI)	> 0.90	Very good (TLI = 0.93)
Root mean square error of approximation (RMSEA)	0.050 < RMSEA < 0.80	Poor (RMSEA = 0.16)
Hoelter's Critical *N*	Hoelter's 0.05 and 0.01 indexes	Good (N = 172 > 68 at 0.05, > 73 at 0.01)
Ration of sample size to number of free parameters	Ratio > 5:1	Very good (ratio 17:1)
3 Fit of internal structure of model		
Individual item reliability	*Pi* > 0.50	Good (COM = 0.77; META = 0.73; SAT = 0.71; EMO = 0.77; ENV = 0.48)
Composite reliability	*Pc* > 0.60	Very good (*pc* = 0.92)
Average variance extracted from a set of measures of a latent variable	> 0.50	Good (0.69)
Significant parameter estimates	Test statistic > or < −1.96 with p-value < 0.05	Very good (All of the parameter estimates were > 10)
Normalized residuals	< 2	Very good (–0.49∼0.79)

internal structure indices. We computed three types of reliability to characterize the Fit of Internal Structure of Model: individual item reliability, composite reliability of the overall scale, and the average variance extracted from the subscales (Bagozzi and Yi 1988) (Table 3). The individual item reliability of the subscales (i.e. squared standardized factor loading) was 0.77 for Commitment Control, 0.73 for Metacognitive Control, 0.71 for Satiation Control, 0.77 for Emotion Control, and 0.48 for Environment Control. The result showed that only the value of Environment Control was slightly less than, but on the verge of, the suggested threshold value of 0.50, whereas the other four measures were well within the satisfactory range. In addition, the overall reliability of the whole SRCvoc scale (i.e. composite reliability) was as high as 0.92,[5] and this was far beyond the recommended value 0.60. Finally, the average variance extracted from these five subscales amounted to 69 per cent,[6] which again exceeded the recommended threshold value of 0.50. These results, together with the reliability information gained from the first study, have suggested that the SRCvoc has good overall reliability.

Finally, an examination of the unstandardized parameter estimates revealed all estimates to be both reasonable and statistically significant at the 0.05 level. That is, all of the parameter estimates were >1.96, indicating that all the parameters were important to the hypothesized model and therefore should be retained in the final construct (Byrne 2001). Also, the fact that all of the normalized residuals were <2 showed that no large amounts of unexplained variance existed and the possibility of a specification error was low (Bagozzi and Yi 1988).

The suggested final construct along with the factor loadings on the paths is presented in Figure 1. An examination of the three types of overall measures

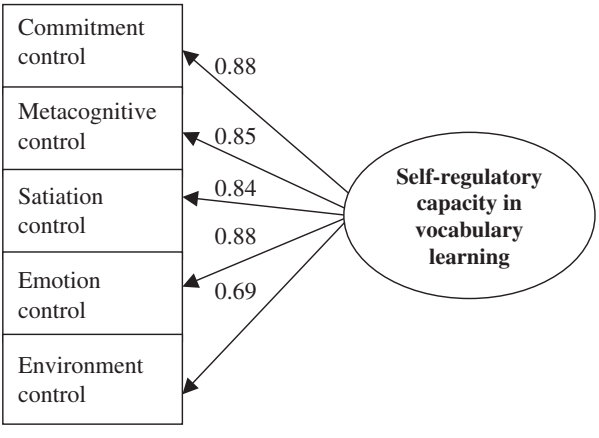

Figure 1: Confirmatory factor analysis of the hypothesized model

of fit revealed a dominant pattern of good support of the overall model as proposed, indicating that the hypothesized construct was empirically valid. Thus, the latent construct of Self-Regulatory Capacity in Vocabulary Learning can be legitimately represented and measured via five indicators: Commitment Control, Metacognitive Control, Satiation Control, Emotion Control, and Environment Control. Accordingly, we believe that the SRCvoc is a meaningful and valid measure and can serve as a basis for exploring the theoretical nature of Self-Regulation.

Exploratory Factor Analysis

Although we have established both the construct reliability and validity of the instrument, it was still not clear whether the scale was unidimensional. That is, it was necessary to determine that no other underlying traits would be uncovered via these five indicators. Support for this claim can be obtained through Exploratory Factor Analysis, and if the unidimensionality of the instrument can be proven then the instrument will be more robust in the sense that it not only has good reliability and validity but also taps into one single underlying trait. We therefore examined whether the five subscales—Commitment Control, Metacognitive Control, Satiation Control, Emotion Control, and Environment Control—loaded primarily on one and the same factor, and if so, how well the underlying construct (i.e. Self-Regulatory Capacity) explained these five indicators.

The five variables were subjected to Principal Axis Factoring.[7] The results revealed that one factor explained the vast majority of the variance (over 69 per cent) and the eigenvalue of the second largest factor was marginal compared to the first one (0.49 vs. 3.75). This provided unambiguous confirmation for the unidimensionality of the measures, and the factor loadings presented in Table 4 also showed a consistently high pattern. We should note, however, that the loading associated with Environment Control was again slightly lower than the rest.

Table 4: Factor loadings on one unrotated factor

	Factor 1
Commitment	0.88
Emotion	0.87
Metacognitive	0.86
Satiation	0.85
Environment	0.69

Extraction Method: Principal Axis Factoring.

CONCLUSION

In this paper we highlighted the definitional fuzziness of language learning strategies as well as the inadequacy of psychometric instruments that have been developed to measure the capacity of strategic learning. In an attempt to offer a solution, we presented a new conceptual approach for operationalizing strategic learning and introduced a new instrument, the 'Self-Regulating Capacity in Vocabulary Learning' scale (SRCvoc), to measure language learner self-regulation in a situated manner. The items of SRCvoc, instead of being operationalized as specific behavioural descriptions, were designed in a way that they would tap into general trends and inclinations, and the theoretical basis of the proposed construct was provided by Dörnyei's (2001) framework of self-regulation. This instrument was then put to the test and a series of elaborate statistical analyses revealed that the newly-conceived operational mechanism of self-regulatory capacity could be empirically validated using vocabulary learning as the agent. The various psychometric indices provided strong evidence that the internal structure of the model was meaningful, appropriate, and thus deserves attention.

While it is important in itself to present a psychometrically sound measure of strategic vocabulary learning, we believe that our results have more general bearings: on the one hand, they indicate that it is possible to design a self-report measure that can assess the degree of a learner's self-regulatory capacity in a given learning domain adequately. On the other hand, the fact that the validity and reliability of SRCvoc are satisfactory and the construct is unidimensional implies that the construct of self-regulation taken from educational psychology can indeed be transferred successfully to the field of second language learning. Our model indicates that self-regulation can be broken into five facets, as outlined by Dörnyei (2001): commitment control, metacognitive control, satiation control, emotion control, and environmental control. Although in our study the construct has been specifically applied in vocabulary research, we believe that the underlying theory and the questionnaire development approach presented above in detail can also be viewed as a heuristic point of departure in the realm of how to devise instruments for use in other learning domains.

Regarding the teaching implications of the model, it can help teachers to understand that the most important aspect of strategic learning is not the exact nature of the specific techniques that students employ but rather the fact that they choose to exert creative effort in trying to improve their own learning. This has been very clearly expressed by Chamot and Rubin (1994) when they stated that 'the good language learner cannot be described in terms of a single set of strategies but rather through the ability to understand and develop a personal set of effective strategies' (p. 372). In other words, the essential aspect of empowering learners is to set into motion the self-regulatory process rather than to offer the instruction of a set of strategies. The latter is undoubtedly a necessary element of the 'learning to learn'

process but it will only be effective if it is supported by an adequate foundation of self-regulatory capacity in the learners.

We suggest that SRCvoc can serve as a diagnostic measure to identify and understand learners' strengths and weaknesses in terms of the five dimensions of self-regulation in the area of (English vocabulary) learning. Of course, SRCvoc cannot provide all the answers because it is an 'off-line' self-report scale and therefore it cannot identify any causes of self-regulatory weakness; as with most quantitative tests and questionnaires in educational psychology, the SRCvoc index can only serve as a starting point for developing appropriate remedies for any self-regulatory shortfalls and researchers need to apply other, more qualitative methodologies (such as stimulated recall and structured observation) to achieve a fuller understanding of the whole picture. However, we hope that the availability of such a diagnostic tool can play an important role in promoting the training of self-regulated learners because, as Randi and Corno (2000) conclude, 'Self-regulation is both an aptitude for and a potential *outcome* of schooling' (p. 651, emphasis added).

Final version received January 2005

APPENDIX 1

Questions asked at the focus-group interviews

1 What do you think of your willpower to achieve the vocabulary learning goals, if any, which you set for yourself?

   (FOLLOW-UPS: Do you feel satisfied/confident about your willpower? Any other special feeling about it? Do you feel the methods to control your willpower are effective/useful? Any other reflections on the methods used?)

2 Think back to an English lesson with many words to learn. Would you feel nervous?

   (FOLLOW-UPS: Do you feel bored/afraid/impatient about this situation? If so, what do you normally do about this feeling? Do you feel satisfied/confident about your methods to control this negative feeling? How effective/useful are the methods?)

3 Suppose that you will have to take an important vocabulary test to pass the course tomorrow. What will you do about this situation?

   (FOLLOW-UPS: Will you procrastinate in reviewing the words? If so, what makes you procrastinate? Comment on the methods used to stop procrastination. Do you feel easily distracted? If so, why do you feel easily distracted? Comment on the methods used to enhance concentration.)

4 Suppose that you are in the process of memorising the new words of a lesson. Will you feel easily bored with this learning activity?

(FOLLOW-UPS: If so, why do you feel easily bored? Comment on the methods used to get rid of the feeling of boredom.)

5 Think back to the days when you had to review words as part of your own study plan, if any. How would you select a conducive time and place for your learning routine?

(FOLLOW-UPS: Would you do this learning activity in a particular context? PROBES: waiting for a bus, during a classroom break, your own room. If so, how important is it to choose a suitable place for studying? Comment on the methods used to control the learning environment.)

APPENDIX 2: 'SELF-REGULATING CAPACITY IN VOCABULARY LEARNING SCALE' (SRCVOC)

English version of the instrument

Dear Fellow Students,

This is an educational research project about learning vocabulary. Below is a series of statements about your learning experience of vocabulary. We would like to know how far these statements match your own perceptions, that is, your personal view. There are no 'right' or 'wrong' answers. Moreover, the data we collect are for research purposes and your opinions will be respected and kept confidential.

There are *twenty items* in total in the questionnaire. Please tick the appropriate box concerning your personal vocabulary learning experience. Thank you very much for your cooperation!

Item	Learning experience	Strongly agree	Agree	Partly agree	Slightly disagree	Disagree	Strongly disagree
1.	Once the novelty of learning vocabulary is gone, I easily become impatient with it.	☐	☐	☐	☐	☐	☐
2.	When I feel stressed about vocabulary learning, I know how to reduce this stress.	☐	☐	☐	☐	☐	☐
3.	When I am studying vocabulary and the learning environment becomes unsuitable, I try to sort out the problem.	☐	☐	☐	☐	☐	☐
4.	When learning vocabulary, I have special techniques to achieve my learning goals.	☐	☐	☐	☐	☐	☐
5.	When learning vocabulary, I have special techniques to keep my concentration focused.	☐	☐	☐	☐	☐	☐
6.	I feel satisfied with the methods I use to reduce the stress of vocabulary learning.	☐	☐	☐	☐	☐	☐
7.	When learning vocabulary, I believe I can achieve my goals more quickly than expected.	☐	☐	☐	☐	☐	☐
8.	During the process of learning vocabulary, I feel satisfied with the ways I eliminate boredom.	☐	☐	☐	☐	☐	☐
9.	When learning vocabulary, I think my methods of controlling my concentration are effective.	☐	☐	☐	☐	☐	☐
10.	When learning vocabulary, I persist until I reach the goals that I make for myself.	☐	☐	☐	☐	☐	☐
11.	When it comes to learning vocabulary, I have my special techniques to prevent procrastination.	☐	☐	☐	☐	☐	☐
12.	When I feel stressed about vocabulary learning, I simply want to give up.	☐	☐	☐	☐	☐	☐
13.	I believe I can overcome all the difficulties related to achieving my vocabulary learning goals.	☐	☐	☐	☐	☐	☐

14.	When learning vocabulary, I know how to arrange the environment to make learning more efficient.	☐	☐	☐	☐	☐	☐
15.	When I feel stressed about my vocabulary learning, I cope with this problem immediately.	☐	☐	☐	☐	☐	☐
16.	When it comes to learning vocabulary, I think my methods of controlling procrastination are effective.	☐	☐	☐	☐	☐	☐
17.	When learning vocabulary, I am aware that the learning environment matters.	☐	☐	☐	☐	☐	☐
18.	During the process of learning vocabulary, I am confident that I can overcome any sense of boredom.	☐	☐	☐	☐	☐	☐
19.	When feeling bored with learning vocabulary, I know how to regulate my mood in order to invigorate the learning process.	☐	☐	☐	☐	☐	☐
20.	When I study vocabulary, I look for a good learning environment.	☐	☐	☐	☐	☐	☐

Note: Commitment control: items 4, 7, 10, 13; metacognative control: items 5, 9, 11, 16; satiation control: items 1, 8, 18, 19; emotion control: items 2, 6, 12, 15; environmental control: items 3, 14, 17, 20.

NOTES

1 We would like to express our sincere gratitude to Dr Peter Y. H. Chen (Department of Foreign Languages and Applied Linguistics, National Taipei University) for his initial comments on the wordings of the Chinese version; Dr Yee-Jean Fon (Graduate Institute of Linguistics, National Taiwan University) for her valuable English–Chinese and Chinese–English translation of the instrument; Dr Ding-Ren Tsao (Department of Foreign Languages and Applied Linguistics, National Taipei University) for her double-checking the first draft of the translation.

2 Amos (Arbuckle and Wothke 1999) is one of the best-known currently available structural equation modelling softwares, also supported by SPSS. Users can specify models via both path diagrams and command lines. A very practical introduction is provided by Byrne (2001).

3 Other indices such as 'absence of negative error variances' and 'error variances not significantly different from zero' are not available in Amos 4.0.

4 Other indices such as 'achievement of adequate statistical power of χ^2 test', 'achievement of linear Q-plot of normalised residual with slope greater than one', and 'high coefficients of determination' are not available in Amos 4.0.

5 Construct reliability

$$= \frac{(\text{Sum of standardized loading})^2}{[(\text{Sum of standardized loading})^2 + \text{Sum of indicator measurement error}]}$$

6 Variance extracted

$$= \frac{\text{Sum of squared standardized loadings}}{[\text{Sum of squared standardized loadings} + \text{Sum of indicator measurement error}]}$$

7 Prior to performing PAF, the appropriateness of data for factor analysis was assessed. The inspection of the correlation matrix revealed that all of the coefficients were above 0.3 (Tabachnick and Fidell 2001); the Kaiser–Meyer–Oklin value was found to be 0.86, exceeding the recommended value of 0.60 (Kaiser 1970, 1974); and the Bartlett's Test of Sphericity (Bartlett 1954) also reached statistical significance ($\chi^2 = 618.34$, $df = 10$, $p < 0.001$), supporting the factorability of the correlation matrix. These three analyses support the suitability of the data for factor analysis.

REFERENCES

Adolphs, S. and **N. Schmitt.** 2003. 'Lexical coverage of spoken discourse,' *Applied Linguistic* 24: 425–38.

Anderson, N. J. 2003. 'Strategy research and training,' Special issue. *TESL-EJ* 7.

Arbuckle, L. J. and W. Wothke. 1999. *Amos 4.0 User's Guide*, Chicago, IL: SmallWaters.

Bagozzi, R. P. and **Y. Yi.** 1988. 'On the evaluation of structural equation models,' *Journal of the Academy of Marketing Science* 16: 74–94.

Bartlett, M. S. 1954. 'A note on the multiplying factors for various chi square approximations,' *Journal of the Royal Statistical Society. Series B* 16: 296–8.

Byrne, B. M. 2001. *Structural Equation Modelling With AMOS: Basic Concepts, Applications, and Programming*. Mahwah, NJ: Lawrence Erlbaum.

Carter, R. 1998. *Vocabulary: Applied Linguistic Perspective* (2nd edn). London: Routledge.

Chamot, A. U. and **J. Rubin.** 1994. 'Comments on Janie Rees-Miller's ''A critical appraisal of learner training: Theoretical bases and teaching implications'' '. *TESOL Quarterly* 28: 771–6.

Chamot, A. U., S. Barnhardt, P. B. El-Dinary, and **J. Robbins.** 1999. *The Learning Strategies Handbook*. New York: Longman.

Cohen, A. D. 1998. *Strategies in Learning and Using a Second Language*. Harlow: Longman.

Cohen, A. D. 2002. 'Preparing teachers for styles- and strategies-based instruction,' in V. Crew, C. Davison, and B. Mak (eds): *Reflecting on Language in Education*. Hong Kong: The Hong Kong Institute of Education.

Corno, L. and **R. Kanfer.** (1993). 'The role of volition in learning and performance,' *Review of Research in Education* 19: 301–41.

DeVellis, R. F. 1991. *Scale Development: Theory and Applications*. London: Sage Publications, Inc.

Dörnyei, Z. 2001. *Motivational Strategies in the Language Classroom*. Cambridge: Cambridge University Press.

Dörnyei, Z. 2003. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. 2005. *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. and **P. Skehan** 2003. 'Individual differences in second language learning,' in C. J. Doughty and M. H. Long (eds): *The Handbook of Second Language Acquisition*. Oxford: Blackwell.

Ehrman, M. E., B. L. Leaver, and **R. L. Oxford.** 2003. 'A brief overview of individual differences in second language learning,' *System* 31: 313–30.

Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Gass, S. M. and **A. Mackey.** 2000. *Stimulated Recall Methodology in Second Language Research*, Mahwah, NJ: Lawrence Erlbaum.

Goulden, R., I. S. P. Nation, and **J. Read.** 1990. 'How large can a receptive vocabulary be?' *Applied Linguistics* 11: 341–463.

Grenfell, M. and **V. Harris.** 1999. *Modern Languages and Learning Strategies: In Theory and Practice*. London: Routledge.

Gu, Y. and **K. R. Johnson.** 1996. 'Vocabulary learning strategies and language learning outcomes,' *Language Learning* 46: 643–79.

Hair, J. F., R. E. Anderson, R. L. Tatham, and **W. C. Black.** 1998. *Multivariate Data Analysis* (5th edn). Upper Saddle River, NJ: Prentice Hall.

Hatch, E. and **A. Lazaraton.** 1991. *The Research Manual*. Boston, MA: Heinle & Heinle.

Kaiser, H. 1970. 'A second generation Little Jiffy,' *Psychometrika* 35: 401–15.

Kaiser, H. 1974. 'An index of factorial simplicity,' *Psychometrika* 39: 31–6.

Kaplan, R. M. and **D. P. Saccuzzo.** 1993. *Psychological Testing: Principles, Applications, and Issues* (5th edn). Belmont, CA: Wadsworth.

Krueger, R. and **M. A. Casey.** 2000. *Focus Groups: A Practical Guide for Applied Research* (3rd edn). London: Sage.

Kuhl, J. 1987. 'Action control: The maintenance of motivational states' in F. Halish and J. Kuhl (eds): *Motivation, Intention, and Volition*. Berlin: Springer.

Kuhl, J. and **T. Goschke.** 1994. 'A theory of action control: Mental subsystems, modes of control, and volitional conflict-resolution strategies' in J. Kuhl and J. Beckmann (eds): *Volition and Personality: Action Versus State Orientation*. Seattle, WA: Hogrefe & Huber.

Lan, R. and **R. L. Oxford.** 2003. 'Language learning strategy profiles of elementary school students in Taiwan,' *IRAL* 41: 339–79.

Laufer, B. 1997. 'The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess' in J. Coady and T. Huckin (eds): *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. New York: Cambridge University Press.

Macaro, E. 2001. *Learning Strategies in Foreign and Second Language Classrooms*. London: Continuum.

MacIntyre, P. D. 1994. 'Toward a social psychological model of strategy use,' *Foreign Language Annals* 27: 185–95.

McDonough, S. 1995. *Strategy and Skill in Learning A Foreign Language*. London: Edward Arnold.

McDonough, S. 1999. 'Learner strategies,' *Language Teaching Research* 32: 1–18.

Nation, I. S. P. 1990. *Teaching and Learning Vocabulary*. Boston, MA: Newbury House.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, I. S. P. and **P. Meara.** 2002. 'Vocabulary' in N. Schmitt (ed.): *An Introduction to Applied Linguistics*. London: Arnold.

Nation, I. S. P. and **R. Waring.** 1997. 'Vocabulary size, text coverage, and word lists' in N. Schmitt and M. McCarthy (eds): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

Nunan, D. 1997. 'Strategy training in the classroom: An empirical investigation,' *RELC Journal* 28: 56–81.

O'Malley, J. M. and **A. U. Chamot.** 1990. *Learning Strategies in Second Language Acquisition*. New York: Cambridge University Press.

O'Malley, J. M., A. U. Chamot, G. Stewner-Manzanares, R. Russo, and **L. Kupper.** 1985. 'Learning strategy applications with students of English as a second language,' *TESOL Quarterly* 19: 557–84.

Oxford, R. L. 1990. *Language Learning Strategies: What Every Teacher Should Know*. New York: Newbury House.

Oxford, R. L. (ed.) 1996. *Language Learning Strategies Around the World: Cross-Cultural Perspectives*. Honolulu, HI: University of Hawaii Press.

Pintrich, P. R., D. A. F. Smith, T. Garcia, and **W. J. McKeachie.** 1991. *A Manual for the Use of The Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, Michigan: The University of Michigan.

Purdie, N. and **R. Oliver.** 1999. 'Language learning strategies used by bilingual school-aged children,' *System* 27: 375–88.

Purpura, J. E. 1999. *Learner Strategy Use and Performance on Language Tests: A Structural Equation Modelling Approach*. Cambridge: Cambridge University Press.

Randi, J. and **L. Corno.** 2000. 'Teacher innovations in self-regulated learning' in M. Boekaerts, P. R. Pintrich, and M. Zeidner (eds): *Handbook of Self-Regulation*. San Diego: Academic Press, pp. 651–85.

Riding, R. and **S. G. Rayner.** 1998. *Cognitive Styles and Learning Strategies: Understanding Style Differences in Learning and Behaviour*. London: David Fulton.

Schmitt, N. 1997. 'Vocabulary learning strategies' in N. Schmitt and M. McCarthy (eds): *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

Schumacker, R. E. and **R. G. Lomax.** 1996. *A Beginner's Guide to Structural Equation Modelling*. Mahwah, NJ: Lawrence Erlbaum.

Skehan, P. 1989. *Individual Differences in Second Language Learning*, London: Edward Arnold.

Stewart, D. W. and **P. N. Shamdasani.** 1990. *Focus Groups: Theory and Practice*. London: Sage.

Stoffer, I. 1995. University foreign language students' choice of vocabulary learning strategies as related to individual difference variables. PhD dissertation. University of Alabama.

Tabachnick, B. G. and **L. S. Fidell.** 2001. *Using Multivariate Statistics* (4th edn). Needham Heights, MA: Allyn & Bacon.

Tseng, W. T. 2000. A Survey of University Students' Vocabulary Knowledge in Taiwan. Unpublished Master Thesis, University of Nottingham.

VanderStoep, S. W. and **P. R. Pintrich.** 2003. *Learning to Learn: The Skill and Will of College Success*. Upper Saddle River, NJ: Prentice Hall.

Weinstein, C. and **R. Mayer.** 1986. 'The teaching of learning strategies,' in M. Wittrock (ed.): *Handbook of Research on Teaching* 3rd edn. New York: Macmillan.

Weinstein, C. E., J. Husman, and **D. R. Dierking.** 2000. 'Self-regulation interventions with a focus on learning strategies,' in M. Boekaerts, P. R. Pintrich, and M. Zeidner (eds): *Handbook of Self-Regulation*. San Diego: Academic Press.

Wenden, A. 1991. *Learner Strategies for Learner Autonomy: Planning and Implementing Learner Training for Language Learners*. Hemel Hempstead: Prentice Hall.

Wenden, A. and **J. Rubin.** 1987. *Learner Strategies in Language Learning*. Hemel Hempstead: Prentice Hall.

Winne, P. H. and **N. E. Perry.** 2000. 'Measuring self-regulated learning' in M. Boekaerts, P. R. Pintrich, and M. Zeidner (eds): *Handbook of Self-Regulation*. San Diego, CA: Academic Press.

Yamamori, K., T. Isoda, T. Hiromori, and **R. L. Oxford.** 2003. 'Using cluster analysis to uncover L2 learner differences in strategy use, will to learn, and achievement over time,' *IRAL* 41: 381–409.

Zeidner, M., M. Boekaerts, and **P. R. Pintrich.** 2000. 'Self-regulation: Directions and challenges for future research,' in M. Boekaerts, P. R. Pintrich, and M. Zeidner (eds): *Handbook of Self-Regulation*. San Diego: Academic Press.