Feature Article

Autonomy as a predictor of English proficiency

Matthew Apple
Nara National College of Technology

Learner autonomy is often regarded as essential to increased English learning (e.g., Benson, 2001; Little, 1991; Cotterall & Crabbe, 1999). However, there have been relatively few quantitative studies that support claims that EFL autonomy influences English proficiency. The present study examines the predictive potential of EFL autonomy on English language learning as measured by the Test of English for International Communication (TOEIC). A sample of 204 engineering students from a Japanese technical college took an 18-item EFL autonomy questionnaire. After verifying construct validity of the EFL autonomy questionnaire items using Rasch item fit analysis (Rasch, 1960), the data were factor analyzed into two distinct constructs. Results from a multiple regression analysis revealed that both factors from the EFL autonomy questionnaire significantly predicted TOEIC scores and accounted for approximately 14% of differences in English achievement.[1]

学習者オートノミー（自律性）はしばしば英語学習の上達に不可欠とされる(e.g., Benson, 2001; Little, 1991; Cotterall & Crabbe, 1999)。しかし、EFL学習者オートノミーが英語運用能力にプラスの影響を与えるという主張をサポートする定量的研究はあまり行われていない。本論はTOEICスコアを指標として、英語習得に対するEFL学習者オートノミーの予測的潜在可能性を検証する。日本の工業高校に在籍する技術系の学生204名に対して18項目のEFLオートノミーアンケート調査を行った。質問項目の構成概念妥当性についてラッシューアイテム分析(Rasch, 1960)を用いて実証した後、要因分析によりデータは２つの構成因子に分析された。更に重回帰分析の結果、EFLオートノミーアンケート調査の両因子について英語能力に

Apple, M. (2011). Autonomy as a predictor of English proficiency. *OnCUE Journal, 4*(3), 191-216. *OnCUE Journal*, ISSN 1882-0220, is available at http://jaltcue-sig.org one year after publication date.

おける予測的差異が認められ、TOEICスコアにおける約14%の違いとして出現した。

Pioneered by Holec (1981), the term "learner autonomy" became a catch phrase in foreign language education in the 1990s (e.g., Benson & Voller, 1997; Dickinson, 1995; Little, 1991). Mainstream North American educational theorists of self-regulation learning (SRL) described autonomy as a process of adjusting goals, strategies, and motivations for general learning (Butler & Winne, 1995; Carver & Scheier, 1990; Winne, 1995a). On other hand, main proponents of foreign language learner autonomy defined it as an "ability" or a "capacity" for self-directed learning. This capacity for organizing one's learning is theorized to include the five steps of (a) setting goals, (b) determining learning content, (c) choosing methods of learning, (d) monitoring learning progress, and (e) reflecting on what has been learned (Cotterall & Crabbe, 1999; Holec, 1981; Little, 1991; Littlewood, 1996).

In addition to managing or understanding the process of language learning, elements of metacognitive behavior, or learning strategies, have been considered an element of learner autonomy (e.g., Cohen, 1998; Oxford, 1989, 1990a, 1990b; Wenden, 1987). Benson (2001) claimed that "learners who achieve proficiency in foreign languages" did so as at least partly as a result of exerting control over learning processes, including the use of self-management strategies to initiate, control, or direct learning processes (p. 65). He refined learner autonomy into the three components of learning management, cognitive processes, and learning content, stating that the components were "clearly interdependent" (p. 50).

The importance of learner autonomy in Japan

Beginning with Rubin's 1975 seminal study of the "good learner" (Rubin, 1975), teachers and researchers in both English as a Second

Language (ESL) and English as a Foreign Language (EFL) contexts have argued that "good language learners" necessarily are autonomous language learners and have extolled the benefits of incorporating autonomous teaching practices, autonomy-encouraging materials, and autonomy-directed learning principles into ESL courses and program curricula (e.g., Cotterall, 2000; Esch, 1996; Little, Ridley, & Ushioda, 2003; Littlewood, 1999; Naiman, 1975, 1996; Nunan, 1988, 1997; Pemberton, Edward, Orr, & Pierson, 1996).

In Japan, the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) revised its educational policies to include the encouragement of EFL autonomy as an educational goal (MEXT, 2003). Educators based in Japan have published several EFL autonomy-related articles and books, including the importance of cultural context (Aoki & Smith, 1999), the relation with cooperative learning (Murphey & Jacobs, 2000), and the use of language learning strategies (Learning Strategy Study Group of Japan Association of College English Teachers, 2005, 2006). A popular Japanese language monthly magazine about English education even devoted an entire issue to EFL autonomy, considering potential benefits for Japanese learners and how to incorporate autonomy into classroom practice (*The English Teachers' Magazine*, February 2008).

Indeed, a common thread among EFL autonomy–related research is a focus on learner-centered classroom activities that promote autonomy. For example, the action research reported in several edited books focused primarily on classroom practices such as the use of portfolios and self-reflective questionnaires (Barfield & Nix, 2003; MacKenzie & McCafferty, 2002; McCasland & Robertson, 2005; Skier & Kohyama, 2006). The classroom-based action research in these publications generally examined ways of encouraging EFL autonomy in the language classroom, with a majority of studies reporting participants for their thoughts and perceptions of learner-centered activities using individual interview or questionnaires. However, few studies listed

the participants' English proficiency level and none demonstrated any measurable improvement of English proficiency as a result of EFL autonomy-based classroom practice.

Remaining questions about learner autonomy

Despite the quantity of publication of articles and books regarding EFL autonomy theory and principles, however, there is still an overall lack of generalizable empirical evidence that autonomy leads to increased foreign language proficiency. Benson (2001) acknowledged the "pressing need for empirical research" to demonstrate a clear connection between EFL autonomy and increased language proficiency (p. 189). In addition to evidence of increased language proficiency due to EFL autonomy, a major aim of such research is the creation and verification of tools for assessing EFL autonomy (Breen & Mann, 1997; Benson, 2001). The vast majority of such tools in the field of SLA are qualitative in nature.

Current available EFL autonomy research primarily consists of interview case studies, diary studies, or learner histories of autonomous learning and strategy usage by the so-called good learner, which typically are limited to a handful of advanced ESL students studying in English-speaking countries (e.g., Benson, 2007; Benson & Nunan, 2005; Cotterall, 2008; Wenden, 1987; Yang, 1998). While case studies and learner histories can shed light on an individual language learner's strategy use and success in language learning, it is difficult to generalize findings from studies with small numbers of highly proficient study participants to larger sample populations of low to medium-proficient students. Unfortunately, there seem to be few existing quantitative studies directly concerning EFL autonomy, particularly for learners in an EFL context, even including studies (Cotterall, 1995, 1996) aimed at learner beliefs about autonomy.

Additionally, there is a lack of empirical evidence demonstrating that autonomous behaviors, such as choosing materials and performing self-evaluation and reflection, are related to or influence English proficiency. There have been a few studies in an English-language, North American

context that demonstrated a connection between autonomy behaviors and beliefs and overall academic performance as measured by course grades (e.g., Cantwell & Moore, 1996; Garcia & Pintrich, 1996); however, there is little convincing empirical evidence that EFL autonomy directly leads to or predicts improved foreign language performance. Finally, there is no empirical evidence that improved EFL autonomy can help Japanese students improve their English proficiency.

The multi-dimensional nature of autonomy has perhaps made quantifiable empirical studies of EFL autonomy difficult to implement. Benson (2001) pointed out that the three learner autonomy components of learning management, cognitive processes, and learning content were "clearly interdependent" (p. 50), and that measuring each individual component of autonomy was difficult due to the "fact that autonomy is clearly a multidimensional construct" (p. 51). While it may be a simple matter of semantics, the issue as discussed by Benson is clearly one of the multidimensional nature of the *concept*, rather than the *construct*, of autonomy, as constructs by definition must be unidimensional (Thurstone, 1931; Wright, 1999). However, given that the concept of language learner autonomy may very well be multidimensional, a plausible first step for measuring learner autonomy therefore is to create measurement tools that demonstrate the unidimensionality of individual constructs of EFL autonomy.

Shimo (2008) first attempted to determine the number of constructs in EFL autonomy for a Japanese learner population by using an 18-item questionnaire designed to measure the multi-construct nature of autonomy according to goal-setting, planning methods of learning, choosing learning materials, and reflecting on language learning. In a study with 106 Japanese university students, she found two factors (Orientation towards improving learning environment and Orientation towards reflective learning), the latter of which significantly correlated (r = .27) with the scores from a mid-term examination using TOEIC-style reading and listening questions (N = 97). In a follow-up study with 148 Japanese

university students (Shimo, 2009), the same 18-item questionnaire was implemented both at the beginning and at the end of a fourteen-week term. Two components were found in the pre-term questionnaire, while for the end of term questionnaire a three-component solution (Enhancing learning opportunities, Reflecting on language abilities, and Planning/ Reflecting on learning processes) was eventually decided upon (Shimo, 2009, p. 41). The question whether the questionnaire contains two or three factors remains open to examination. The answer may be related to construct validity, the demonstration of which may improve the external validity and generalizability of existing study results.

Purpose of the Study

The purpose of the present study was not only to validate constructs present in the EFL autonomy questionnaire (Shimo, 2008, 2009), but also to examine the degree to which EFL autonomy can predict English language learning proficiency for a Japanese student population measured by a widely-used standardized exam. Construct validation was examined using Rasch item fit analysis (Rasch, 1960), which assisted in questionnaire revision to provide more consistent results across sample populations. The ability of EFL autonomy to predict English proficiency was examined by factor analyzing the questionnaire data and using the resultant EFL autonomy factor scores as predictors of scores from the Test of English for International Communication (TOEIC) in a multiple regression analysis.

Methods

Participants

Students from a technical college (kousen) participated in this study. The total participant *N*-size was 204; however, data from 19 participants were removed due to incomplete, illegible, or set response patterns (i.e., answering all questions with the same response) that would adversely affect the correlational matrixes in later analyses. The

resulting *N*-size was 185, comprising 164 third-year students (high school senior equivalent) and 21 first-year students from the Faculty of Advanced Engineering (*senkoka*, university third year equivalent). Third-year students were taking one 90-minute four-skills English class and one 45-minute conversational English class per week in a 30-week academic year, for a total of 67.5 English classroom hours. First-year senkoka students were taking one 90-minute four-skills English class for 30 weeks, and one 90-minute technical English writing course for 15 weeks, for a total of 90 English class hours.

Measurement instruments

Two instruments were used in this study. The first was the EFL autonomy questionnaire created by Shimo (2008), which consisted of 18 items (see Appendix). The questionnaire was implemented during class at the end of May, in the week following the spring mid-term examination. Students were informed that participation was voluntary and that responses to the questionnaire would not influence their course grades in any way.

The second instrument used was the TOEIC-IP exam. The exam was given to all students in the study during the first week of April and the results were returned to them in class at the time of questionnaire implementation. The combined TOEIC scores were used as self-reported by participants, $M = 315.5$, $SD = 71.72$. Although TOEIC has been criticized for its error margin (see the Discussion section), TOEIC was used to measure students' English proficiency for two reasons. First, TOEIC is by far the most widely-used English proficiency exam in Japan, with over 940,000 students sitting for the exam in 2008 to go along with 778,000 general test takers (Transition in TOEIC Test-taker Numbers, n.d.). Second, as a standardized exam of English proficiency, TOEIC provides a greater degree of external generalizability than in-house exams or course grades.

Analysis Techniques

The study used four techniques of data analysis: (a) Rasch measurement analysis, (b) parallel analysis, (c) exploratory factor analysis, and (d) multiple regression analysis. Rasch analysis was used to examine the quality of the individual questionnaire items. Parallel analysis and exploratory factor analysis were used in conjunction to determine the number of factors represented by the data. Multiple regression analysis was used to determine the relationship between factors of EFL autonomy and English proficiency as measured by TOEIC.

Rasch analysis. Data from the EFL autonomy questionnaire was first analyzed according to the Rasch model using WINSTEPS 3.63 software (Linacre, 2006) to check for construct validity. The Rasch model (Rasch, 1960) mathematically transforms raw data scores into logarithms and then places both item responses and persons on the same log-odds scale for the purposes of determining construct unidimensionality and item difficulty (Bond & Fox, 2007).

The primary method of checking for measurement construct validity was the examination of item fit, which refers to how well each item measures the intended construct. When measuring related constructs, item fit has been demonstrated to be a more accurate estimation of construct validity than Cronbach's alpha, which is highly dependent upon the number redundant items and is largely irrelevant to dimensionality or construct validity (Cortina, 1993; Embretson & Reise, 2000; Green, Lissitz, & Mulait, 1977; Schmitt, 1996; Smith, E., 2002; Smith, R., 2000; Sijtmas, 2009). Moreover, item fit statistics can also indicate item and person responses that misfit the model as a result of carelessness, response set answering, or item bias (Wolfe & Smith, 2007b, p. 211).

Item fit was measured by infit and outfit statistics, which have both an unstandardized form and a standardized form. Infit and outfit mean squared fit statistics are chi-square ratios that determine how

well the observed data fit the probabilistic expectations of the Rasch model (Bond & Fox, 2007, p. 238). The mean squared statistics can also be transformed into standardized *z*-scores, which approximate a normalized distribution (Bond & Fox, 2007, p. 57). Fit statistics thus improve generalizability of questionnaire study results across samples (Wilson, 2005; Wolfe & Smith, 2007a; Wright, 1996).

Based on the results in Shimo (2008, pp. 166-167), two sets of items were examined for item fit to their expected constructs. Items 1, 4, 5, 8, 10, 11, 14, and 16 were examined for fit to the intended construct Orientation towards improving learning environment, while items 2, 3, 6, 7, 9, 12, 13, 15, 17, and 18 were examined for fit to the intended construct Orientation towards reflective learning. The infit statistic focuses on the performance of participants whose responses are closest to the endorsability difficulty of the items, and thus infit is typically considered more important than outfit (Bond & Fox, 2007, p. 57). Therefore in this study, an infit mean squared statistic of between .75 and 1.3 and an infit standardized z-score of between -2.0 and 2.0 were used as the criterion for good item fit to the construct (Bond & Fox, 2007, p. 240), with a mean squared fit statistic of 1.0 and a standardized *z*-score of 0 indicating ideal fit to the proposed construct model. Misfitting items were removed prior to factor analysis, since such items add error to the measurement and thus reduce generalizability of the results.

Monte Carlo parallel analysis and factor analysis. A parallel analysis was run to determine the number of factors present in the data (Horn, 1965). The Monte Carlo PCA for Parallel Analysis program (Watkins, 2000) generated a list of random eigenvalues based on the number of questionnaire items and the number of participants. Factors extracted from the data set obtained from the participants were compared to the random eigenvalues from the parallel analysis to assist in determining how many factors to retain. Following parallel analysis, the data were

factor analyzed using principal components analysis in SPSS 15 to confirm the number of factors present in the data and to determine whether the items loaded onto their hypothesized constructs.

The sample size for factor analysis was N = 185, with 13 item variables. Although Tabachnick and Fidell (2007, p. 613) recommended an N-size of 300 cases as "comfortable," they also noted that strong factor loadings might require as few as 150 cases. Stevens (1996) suggested a minimum of five cases per variable, making N = 65 the minimum sample size required for this study. Guadagnoli and Velicer (1988) additionally indicated that factors consisting of four or more loadings above .60 were reliable regardless of the sample size. Finally, Stevens (2009) noted that reliable factors may be indicated by an "average of the four largest loadings" above .60 (p. 333) and suggested that the most important determinations for factor analysis reliability were the factor loading strengths and the number of variables per factor. Thus, reliability of the factor analysis for the sample size could only be determined following examination of the resultant item factor loadings.

Multiple regression analysis. Factor scores for items on each factor of the foreign language learner autonomy questionnaire were used as predictor variables of the outcome variable, combined TOEIC scores, to determine the effect of learner autonomy on English learning for the sample population. Stevens (2009, p. 117) recommended 15 cases per independent variable, while Tabachnick and Fidell (2007, p. 123) recommended the formula $N \geq 50 + 8m$, where m is the number of independent variables. Thus, based on each of these recommendations, the sample size (N = 185) was more than adequate for multiple regression analysis.

Results

Rasch Analysis

Item fit analysis statistics indicated that five items out of 18 did

not fit the model for each hypothesized construct (see Appendix). Fit statistics for items 4 and 14 on the intended construct Orientation towards improving learning environment, and for items 3, 9, and 17 on the intended construct Orientation towards reflective learning failed to fit the model. A closer look at the wording of the five items revealed that all could have been perceived by participants as being concerned with abilities and testing rather than with aspects of learner autonomy. These items were eliminated from further analysis, leaving a total of 13 questionnaire items. The remaining items were examined again for item fit analysis, and measures were obtained to show the relative strength of the item based on distributed *t*-scores (Table 1). The construct "Orientation towards improving learning environment" had a Rasch item reliability of .97, and the construct of "Orientation towards reflective learning" had a Rasch item reliability of .98.

Table 1. Item Fit Statistics for Orientation Toward Improving Learning Environment and Orientation Towards Reflective Learning

Item	Measure	Infit Mean Square	Infit *z*-score	Outfit Mean Square	Outfit *z*-score
Orientation Toward Improving Learning Environment					
15	44.43	1.28	2.5	1.40	3.3
5	52.79	1.28	2.3	1.22	1.8
18	47.46	.92	-.7	1.00	.0
11	60.01	.99	-.1	.82	-1.4
16	51.03	.95	-.4	.88	-1.1
8	43.10	.83	-1.6	.89	-1.0
1	51.19	.82	-1.7	.75	-2.3
Orientation Towards Reflective Learning					
12	47.19	1.16	1.5	1.17	1.6
6	37.66	1.13	1.3	1.13	1.2
13	49.29	1.06	.6	1.04	.4
2	54.00	.93	-.7	.94	-.5
7	56.95	.85	-1.5	.84	-1.5
10	54.91	.81	-1.9	.83	-1.7

Notes. "Measure" = item difficulty as measured by *t*-scores, with a mean of 50; *N* = 185.

Table 2. Comparison of Eigenvalues from a Randomly-generated Data Set and the Sample Data

Factor number	Actual Eigenvalue from sample	Random value from parallel analysis	Decision
1	6.33	1.46	Accept
2	1.40	1.35	Accept
3	.83	1.25	Reject

Monte Carlo Parallel Analysis and Factor Analysis

To verify the number of factors to be extracted from the data set, a Monte Carlo Parallel Analysis was run (Watkins, 2000). Random eigenvalues generated from the parallel analysis compared to results from the factor analysis of data in the study indicated the presence of two factors (Table 2).

Following confirmation of the number of factors, factor analysis was conducted (Table 3). The KMO measure for sampling adequacy was .905, and the Bartlett's Test of Sphericity was significant at .000, supporting the factorability of the correlation matrix. A varimax rotation was used to assist in interpreting the results, using a minimum of .40 for factor loadings (Stevens, 2009, p. 333). The first factor consisted of seven items ranging in factor loadings from .57 to .80, while the second factor consisted of six items ranging in factor loadings from .61 to .81. Two items, item 2 and item 10, loaded onto both factors at above the .40 cut-off point. Total variance accounted for by both factors was 59.4%. The strength of the factor loadings exceeded the minimum recommended for determining reliable factors (Guadagnoli & Velicer, 1988; Stevens, 2009); thus, assumptions for the factor analysis were met and the results should be considered valid for this sample.

Multiple Regression Analysis

Factor scores from the EFL autonomy questionnaire items were correlated with combined TOEIC reading and listening scores to

Table 3. Factor Analysis of Foreign Language Learning Autonomy for a Japanese University Sample

Item	Item description	Factor 1	Factor 2	h^2
8	I usually try to be in touch with English outside class	**.80**	.23	.69
1	I am usually willing to read in English; for example, I read English newspapers and books	**.76**	.30	.61
5	I am usually willing to watch English TV programs and/or listen to English radio programs	**.74**	.07	.55
16	I usually try to seek opportunities to speak in English	**.73**	.24	.58
11	I am usually willing to write in English; for example, I write email or keep a diary in English	**.70**	.22	.53
18	I usually study English by using learning materials of my own choice	**.68**	.39	.62
15	I usually think whether my English learning materials and dictionaries are useful or not	**.57**	.32	.43
12	I usually think why I did so when I made mistakes in English exercises	.09	**.81**	.66
6	I usually check what I don't understand in my English homework in some way	.18	**.71**	.54
7	I usually plan my English studying	.29	**.71**	.58
13	I usually think whether my way of studying English is effective for me or not	.26	**.71**	.57
2	I usually have a plan about how to study English	.44	**.66**	.63
10	I usually set goals on my own about studying English	.54	**.61**	.66

Notes. Principal component analysis with varimax rotation was used. F1= "Orientation toward improving language learning environment" (α = .88); F2= "Orientation towards reflective learning" (α = .86); h^2 = commonalities; *N* = 185.

Table 4. EFL Autonomy Predictors for the Independent Variable TOEIC Scores

Factor	*B*	*SEB*	*β*	*t*
Orientation towards improving learning environment	.335	.069	.335	4.877*
Orientation towards reflective learning	.176	.069	.176	2.562*

Notes. R^2 = .14; Adjusted R^2 = .13; N = 185.
*p < .01

determine the feasibility of a multiple regression analysis. TOEIC scores correlated significantly (p < .01) with both *Orientation towards improving learning environment* (r = .34) and *Orientation towards reflective learning* (r = .18). The two EFL autonomy factors correlated significantly with each other (r = .64). Although Tabachnick and Fidell (2007, p. 125) indicated that correlations as high as .6 may pose problems in interpreting multiple regression analysis results, examination of the regression residual plots revealed that the assumption of lack of multicollinearity had been met. Checking the Mahalanobis distances and Cook's Distance also verified that no outliers were present in the data. Thus, the assumptions of multicollinearity, homoscedasticity, outliers, and independence of residuals among the two factors from the EFL autonomy questionnaire and the combined TOEIC scores were checked and met for multiple regression analysis.

A hierarchical regression analysis was performed, with *Orientation towards improving learning environment* and *Orientation towards reflective learning* as independent predictor variables, and TOEIC scores as the dependent outcome variable (Table 4). Results indicated that both EFL autonomy factors predicted TOEIC scores, $F(2,181) = 15.176$, R^2 = .14, p = .000, with EFL autonomy accounting for 14% of the variances in TOEIC scores. The change in R^2 indicated that the construct *Orientation towards improving learning environment* alone accounted for 11% of the differences in TOEIC scores among sample participants.

Discussion

The first conclusion to be drawn from the results is the stability of the constructs represented by the questionnaire items. Five of the original 18 items—3, 4, 9, 14, and 17—were found to misfit their intended constructs in the Rasch analysis. Two of these items (9 and 17) were similarly removed in Shimo (2009) on the basis of low commonality values during factor analysis. Removal of the five misfitting items resulted in two unidimensional constructs with Rasch item reliabilities of .97 and .98. Parallel analysis and factor analysis confirmed the presence of two independent, but related, constructs in the data, with the two constructs accounting for nearly 60% of the variance. Therefore, it is reasonable to expect similar results from future studies using the EFL autonomy questionnaire in this study.

The second conclusion to be drawn concerns the degree to which EFL autonomy can predict English learning outcomes. Among the study participants, 14% of the differences in TOEIC scores were predicted by EFL autonomy variables. Furthermore, 11% was predicted only by the construct *Orientation towards improving learning environment*. In other words, students who actively sought English language learning materials in addition to the regular classroom materials tended to have higher combined TOEIC scores.

The multiple regression results for *Orientation toward improving learning environment* can also answer the question of the degree to which EFL autonomy leads to an increase in English proficiency. The standardized beta weight scores (β) of predictor factors in multiple regression analysis are directly related to the standard deviation of the outcome variable. Thus, if students' responses to EFL autonomy items were to change by one standard deviation unit, their TOEIC scores would also change by one standard deviation unit. For the sample in this study, the TOEIC standard deviation was 71.72; multiplying .335, the beta of the construct *Orientation towards improving learning environment*, times the TOEIC standard deviation revealed that an increase in

students' orientation towards improving learning environment (i.e., actively searching and discovering English learning materials outside the classroom) of one standard deviation would predict an increase of roughly 24 points on the TOEIC.

On the other hand, the standard error of measurement of TOEIC is 25 points for each section of the test, and the error between two different test implementations may be as high as 35 points (*TOEIC Examinee Handbook*, 2008, pp. 19-20). Thus, it is difficult to prove with 100% degree of certainty whether the difference in TOEIC scores for the study sample was caused by the presence of EFL autonomy. An additional point regarding the use of TOEIC as a measure of English proficiency is the overall passive nature of the test items. Although an initial validity study sponsored by ETS indicated medium to strong correlations among various subsections of TOEIC and speaking and writing measurements (Woodfard, 1982), results from an examination of EFL autonomy and a direct measurement of "active" English proficiency skills such as an oral interview or essay writing may differ from the current study.

Regardless of the statistical validity problems associated with TOEIC as a measurement instrument, the use of TOEIC to demonstrate English proficiency has legitimate face validity for students in this study. Many students at technical colleges transfer to four-year universities at the end of the five years of the main program via university *hennyuu* (transfer entrance) examinations, many of which use TOEIC scores as a major measurement, if not the sole measurement, of students' English abilities. Students who wish to obtain employment as a technician at a Japanese company are likewise expected to submit TOEIC scores, which can sometimes determine the outcome of job interviews. Therefore, the ability to find language learning materials in order to offset decreased English class contact hours can literally affect the future careers of the engineering students in this study.

Conclusion

The results of this study indicated that the participants with the autonomous language learning habit of seeking English learning materials outside the regular class texts tended to have greater English proficiency than their peers who did not have seek self-selected materials outside class. The participants had minimal classroom exposure to English during a typical week, with two hours and fifteen minutes for the third-year students and three hours for the first-year senkoka students. Increasing exposure to English learning materials was therefore particularly relevant for students in this study given the limited amount of study time available to them.

The current study has contributed to the field of foreign language learner autonomy research by answering the call of Benson (2001) for the demonstration of a connection between learner autonomy and English proficiency. The EFL autonomy questionnaire by Shimo (2008) was found, after revisions, to predict English proficiency as measured by the TOEIC.

However, limitations to the study raise further questions. The participants were all engineering majors, with a relatively low proficiency level of English as measured by TOEIC, who had limited English language classroom time. Therefore, reasonable questions to examine in further research studies include:

- To what degree does EFL autonomy vary among students majoring in different academic fields?
- To what degree is the amount of weekly English classroom hours related to EFL autonomy?
- Is EFL autonomy more or less effective at increasing English proficiency for low-, intermediate-, or high-level proficiency students?
- Would EFL autonomy predict a greater amount of English proficiency as measured by other forms of standardized English assessment (e.g., TOEFL scores or Cambridge Proficiency Exam

levels)?

- Does EFL autonomy lead to increased English proficiency for all language skills or only for some? Which ones and how much?

Future research should examine these and other related issues concerning EFL autonomy in order to provide empirical evidence of its effectiveness over traditional teacher-centered approaches to EFL classroom instruction. Further iterations with larger samples of students in different learning contexts may verify the EFL learner autonomy questionnaire across samples consisting of different study majors, in different educational contexts, and at different levels of English proficiency.

Note

[1]Preliminary results from this study were presented at the JALT CUE 2009 Conference at Tezukayama University, on October 17, 2009.

References

Aoki, N., & Smith, R. (1999). Autonomy in cultural context: The case of Japan. In S. Cotterall & D. Crabbe (Eds.), *Learner autonomy in language learning: Defining the field and effecting change* (pp. 19-28). Frankfurt am Main, Germany: Peter Lang.

Barfield, A., & Nix, M. (Eds.). (2003). *Learner and teacher autonomy in Japan 1: Autonomy you ask.* Tokyo: Japan Association for Language Teaching.

Benson, P. (2001). *Teaching and researching autonomy in language learning*. London: Pearson Education.

Benson, P. (Ed.). (2007). *Learner autonomy 8: Teacher and learner perspectives*. Dublin, Ireland: Authentik.

Benson, P., & Nunan, D. (Eds.). (2005). *Learners' stories: Difference and diversity in language learning.* Cambridge: Cambridge University Press.

Benson, P., & Voller, P. (Eds.). (1997). *Autonomy and independence in language learning.* Harlow: Longman.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences (2nd Ed.).* Mahwah, New Jersey: Lawrence Erlbaum.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245-281.

Cantwell, R. H., & Moore, P. J. (1996). The development of measures of individual differences in self-regulatory control and their relationship to academic performance. *Contemporary Educational Psychology, 21,* 500-517.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97,* 19-35.

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

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

Cotterall, S. (1995). Readiness for autonomy: Investigating learner beliefs. *System, 23*(2), 195-205.

Cotterall, S. (1999). Key variables in language learning: What do learners believe about them? *System, 27*, 493-513.

Cotterall, S. (2000). Promoting learner autonomy through the curriculum: Principles for designing language courses. *English Language Teaching Journal, 54*(2), 109-117.

Cotterall, S. (2008). Autonomy and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 110-120). Cambridge: Cambridge University Press.

Cotterall, S., & Crabbe, D. (Eds.). (1999). *Learner autonomy in language learning: Defining the field and effecting change*.

Frankfurt am Main, Germany: Peter Lang.

Dickinson, L. (1995). Autonomy and motivation: A literature review. *System, 23*(2), 165-174.

The English Teachers' Magazine [Eigokyouiku], *56*(12). (February, 2008). Tokyo: Taishukan.

Garcia, T., & Pintrich, P. R. (1996). The effects of autonomy on motivation and performance in the college classroom. *Contemporary Educational Psychology, 21,* 477-486.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test undimensionality. *Educational and Psychological Measurement, 37*, 827-838.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample sizes to the stability of component patterns. *Psychological Bulletin, 103,* 265-275.

Holec, H. (1981). *Autonomy and foreign language learning.* Oxford: Pergamon.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 20,* 179-185.

Learning Strategy Study Group of Japan Association of College English Teachers [Daigaku Eigo Kyouiku Gakkai Gakushu Sutorateji Kekyukai]. (2005). *Learning strategy handbook for English teachers* [Eigo kyoushi no tameno gakushu sutorateji handobukku]. Tokyo: Taishukan.

Learning Strategy Study Group of Japan Association of College English Teachers [Daigaku Eigo Kyouiku Gakkai Gakushu Sutorateji Kekyukai]. (2006). *Language learning and learning strategies: An applied linguistics approach to autonomous learning* [Gengogakushu to gakushu sutorateji: Jiritsu gakushu ni muketa ouyougengogaku karano apurouchi]. Tokyo: Taishukan.

Little, D. (1991). *Learner autonomy 1: Definitions, issues, and problems.* Dublin, Ireland: Authentik.

Little, D., Ridley, J., & Ushioda, E. (Eds.). (2003). *Learner autonomy in*

the foreign language classroom: Teacher, learner, curriculum, and assessment. Dublin, Ireland: Authentik.

Littlewood, W. (1996). "Autonomy": An anatomy and a framework. *System, 24*(4), 427-435.

Littlewood, W. (1999). Defining and developing autonomy in East Asian contexts. *Applied Linguistics, 20*(1), 71-94.

Linacre, J. M. (2006). WINSTEPS Rasch measurement computer program [Computer software]. Chicago.

MacKenzie, A., & MacCafferty, I. (Eds.). (2002). *Developing autonomy*. Tokyo: JALT Publications.

McCasland, P., & Robertson, M. (Eds.). (2005). *Learner development: Context, Curricula, Content.* Tokyo: JALT Publications.

MEXT (The Ministry of Education, Culture, Sports, Science, and Technology). (2003). Regarding the establishment of an action plan to cultivate "Japanese with English abilities." Retrieved from http://www.mext.go.jp/english/topics/03072801.htm

Murphey, T., & Jacobs, G. M. (2000). Encouraging critical collaborative autonomy. *JALT Journal, 22*(2), 229-244.

Naiman, N., Frolich, M., & Stern, H. H. (1975). *The good language learner.* Toronto: The Ontario Institute for Studies in Education.

Naiman, N. (1996). *The good language learner.* Clevedon: Multilingual Matters.

Nunan, D. (1988). *The learner-centred curriculum: A study in second language teaching.* Cambridge: Cambridge University Press.

Nunan, D. (1997). Designing and adapting materials to encourage learner autonomy. In P. Benson & P. Voller (Eds.), *Autonomy & independence in language learning* (pp. 192-203). London: Longman.

Oxford, R. L. (1989). Use of language learning strategies: A synthesis of studies with implications for strategy training. *System, 17*, 235-247.

Oxford, R. L. (1990a). Language learning strategies and beyond:

A look at strategies in the context of styles. In S. S. Magnan (Ed.), *Shifting the instructional focus to the learner* (pp. 35-55). Middlebury, VT: Northeast Conference on the Teaching of Foreign Languages.

Oxford, R. L. (1990b). *Language learning strategies: What every teacher should know*. New York: Newbury House.

Pemberton, R., Edward, S. L., Orr, W. W. F., & Pierson, H. D. (Eds.). (1996). *Taking control: Autonomy in language learning*. Hong Kong: Hong Kong University Press.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago.

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

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8,* 350-353.

Shimo, E. (2008). Learner autonomy and English language proficiency: An exploration among non-English majors at a Japanese university. *Kinki University Language Education Department Bulletin, 8*(2), 153-178.

Shimo, E. (2009). Dynamics of learner autonomy in language learning: Components and changes. *JACET Kansai Journal, 11,* 33-49.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107-120.

Skier, E., & Kohyama, M. (Eds.). (2006). *Learner and teacher autonomy in Japan 2: More autonomy you ask*. Tokyo: Japan Association for Language Teaching.

Smith, E. V. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *Journal of Applied Measurement, 3*(2), 205-231.

Smith, R. M. (2000). Fit analysis in latent trait measurement models.

Journal of Applied Measurement, 1(2), 199-218.

Stevens, J. P. (1996). *Applied multivariate statistics for the social sciences (3rd Ed.)*. Mahwah, NJ: Erlbaum.

Stevens, J. P. (2009). *Applied multivariate statistics for the social sciences (5th Ed.)*. New York: Routledge.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics (5th Ed.)*. Boston: Pearson Education.

Thurstone, L. L. (1931). Measurement of social attitudes. *Journal of Abnormal and Social Psychology, 26,* 249-269.

TOEIC Examinee Handbook. (2008). Retrieved from http://www.ets.org/toeic

Transition in TOEIC Test-taker Numbers [TOEIC jukenshasuu no suii]. (n.d.). Retrieved from http://www.toeic.or.jp/toeic/pdf/about/transition1979_2008.pdf

Watkins, M. W. (2000). Monte Carlo PCA for parallel analysis [Computer software]. State College, PA: Ed & Psych Associates.

Wenden, A. (1987). Incorporating learner training in the classroom. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 159-178). Englewood Cliffs, NJ: Prentice-Hall.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum.

Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist, 30*(4), 173-187.

Wolfe, E. W., & Smith, E. V. (2007a). Instrument development tools and activities for measure validation using Rasch models: Part I — instrument development tools. *Journal of Applied Measurement, 8*(1), 97-123.

Wolfe, E. W., & Smith, E. V. (2007b). Instrument development tools and activities for measure validation using Rasch models: Part II —validation activities. *Journal of Applied Measurement, 8*(2), 204-234.

Woodfard, P. S. (1982). The Test of English for International

Communication (TOEIC). In C. Brumfit (Ed.), *English for International Communication* (pp. 61-72). New York: Pergamon Press.

Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling, 3*(1), 3-24.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know* (pp. 65-104). Mahwah, NJ: Lawrence Erlbaum.

Yang, N.-D. (1998). Exploring a new role for teachers: Promoting learner autonomy. *System, 26*(1), 127-135.

Matthew Apple *teaches English composition and academic presentation at Nara National College of Technology. His research interests include individual differences in EFL theory and practice, CALL, and educational statistics.* ***mapple@libe.nara-k.ac.jp***

Appendix

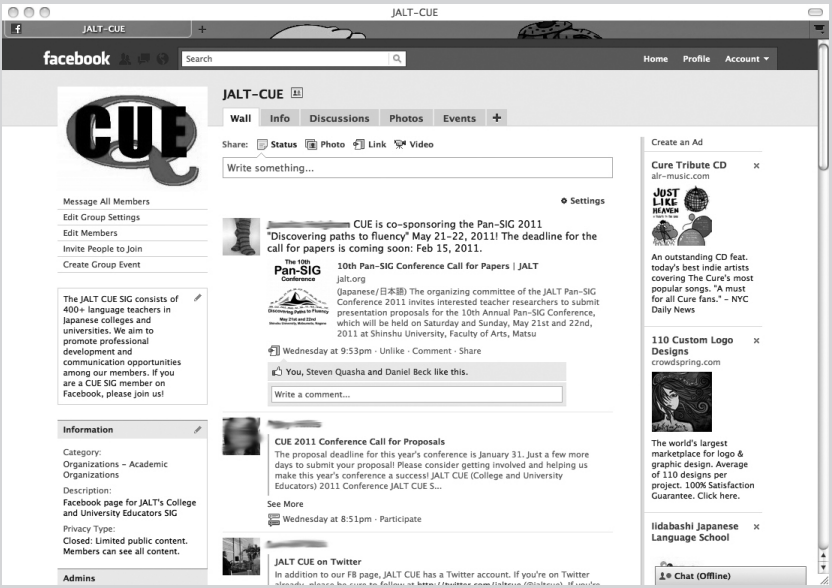
EFL Learner Autonomy questionnaire (Shimo, 2008, 2009)

1. I am usually willing to read in English; for example, I read English newspapers and books.
2. I usually have a plan about how to study English.
3. *I usually look back at my mistakes in English quizzes.
4. *I usually think about what kinds of English skills I want to improve.
5. I am usually willing to watch English TV programs and/or listen to English radio programs.
6. I usually check what I don't understand in my English homework in some way.
7. I usually plan my English studying.
8. I usually try to be in touch with English outside class.
9. *I usually understand which English skill (for example, reading, grammar, listening, etc.) I am poor at.
10. I usually set goals on my own about studying English.
11. I am usually willing to write in English; for example, I write email or keep a diary in English.
12. I usually think why I did so when I made mistakes in English exercises.
13. I usually think whether my way of studying English is effective for me or not.
14. *I am usually willing to take tests such as TOEIC in order to know my own English proficiency.
15. I usually think whether my English learning materials and dictionaries are useful or not
16. I usually try to seek opportunities to speak in English.
17. *I usually understand which English skill (for example, reading, grammar, listening, etc.) I am good at.

18. I usually study English by using learning materials of my own choice.

* Identified as a misfitting item in Rasch analysis and removed prior to factor analysis and multiple regression analysis.

Join us on the JALT-CUE Facebook site to discuss this issue's topics with other CUE members!



On Facebook, search for "**JALT-CUE.**"