

# STUDENT EVALUATIONS AS EFFECTIVE INDICATORS OF LEARNING OUTCOMES: EVIDENCE FROM JAPANESE UNIVERSITY STUDENTS IN ENGLISH LANGUAGE COURSES

EMIKO KOBAYASHI  
JOHN BINTLIFF

## ABSTRACT

The term “student learning outcomes” refers to the knowledge, skills, and abilities that students achieve during a course, and is typically assessed based on student evaluations conducted at the end of the semester. Previous studies in this area have investigated the effects of instructional quality and academic demands separately and have been limited primarily to examining findings using student samples from the United States. With Japanese college students’ perceptions of self-improvement in English language courses as the dependent variable, the present study directly tests the hypothesis that students who perceive instructional quality to be higher, and course demands to be greater, also estimate higher levels of self-improvement in English language skills. The analysis provides strong support for this hypothesis.

In the literature on student evaluation of teaching (SETs), student learning often has been assessed based upon evaluations that students complete regarding the quality of instruction received and the academic demands placed upon them. Recent research finds a link between academic achievement and various indicators of students’ attitudes toward instructional quality, such as clear explanations of subject content and appropriate pacing of instruction (Greimel-Fuhrmann and Geyer 2003), as well

as course demands such as how much time students spend on studying outside the classroom for preparation and homework completion (Nois and Hudson 2006; McFadden and Dart 1992). Research testing the relative effects of students' attitudes toward instructional quality and academic demands on student learning is rare, however, and what has been done so far has largely been restricted to a sample of students in North America, especially in the United States.

Marsh (1987) observed decades ago that most of the research findings on the relationship between student ratings and student learning appear generalizable across different institutions, students, and courses (see also Braskamp and Ory 1994; Centra 1993) – an observation based primarily on the lack of evidence for the relationship outside the U.S. Thus, what we know about student evaluations of course instruction and its relationship to course demands and academic achievement may not be generalizable to students in other countries. Ory and Ryan (2001: 36) note that, “There is some uncertainty about the existence of the relationship between student ratings and student achievement in all contexts.” (See also Brandenburg, Slinde, and Bastista 1977; Cashin 1990; Costin, Greenough, and Menges 1971; Feldman 1978). It is possible that the substantial effects of student evaluations of instruction and course demands reported in the U.S. and other North American samples may not be as evident in other countries (see Marsh 1987 for a discussion).

Higher ratings of instruction and academic demands have already been shown to increase levels of student learning (Greimel-Fuhrmann and Geyer 2003; Nois and Hudson 2006; McFadden and Dart 1992). The present study is the first to provide direct evidence of the relative importance of student evaluations of instructional quality and academic demands as predictors of student learning and the first ever to do so with a sample of Japanese college students enrolled in a required English as a foreign language course. Our hypothesis is that Japanese students who perceive instructional quality to be higher, and course demands to be greater, estimate higher levels of self-improvement in English language skills. Thus we test Japanese students' attitudes toward instructional quality and course demands as

independent variables predicting their perceptions of self-improvement in English language courses. The research focuses on Japanese students' improvement in English language skills because English education in Japan is an arena in which the debate over limited English proficiency rages on, and because other research suggests reconsideration of English education in light of the demands of the rapidly expanding global era (Amaki 2008).

## LITERATURE REVIEW

This research into how attitudes toward instructional quality and time spent studying outside the classroom affect Japanese university students' perceptions of self-improvement in English language courses adds another branch to the already dense but continually thickening forest of work done on Student Evaluation of Teachers.

Research into student assessment dates back to the early twentieth century (Brandenburg and Remmers 1927; Remmers 1927, 1928, 1930; Remmers and Brandenburg 1927) and continues to expand in scope and significance. SETs are not only a critical tool for improving student learning, but they also significantly impact the careers of instructors (Clayson 2009; Ory and Ryan 2001). Seldin (1993), in his extensive research into faculty evaluation, reported that teaching is the number one factor used in assessing overall faculty performance. The fact that there are already over 2,000 published studies dealing with student evaluation of teaching (Murray 2005) further underscores their impact on the field of education.

Over the last eighty years much effort has been devoted to assessing the influence that attitudes toward instructional quality have on student learning, but the research has presented mixed results. Perhaps most famous among studies finding a negative correlation between student rating and learning is Rodin and Rodin (1972). The negative correlation of  $-0.75$  indicates that students tended to give lower ratings to instructors from whom they learned the most, but gave higher evaluations to those

from whom they learned the least. While this study is often cited to discredit the validity of SETs, its findings have subsequently undergone increasingly negative scrutiny.

Conversely, evidence abounds suggesting that instructional quality plays an important role in students' acquisition of knowledge, skills, and understanding (e.g., Brophy 1986/1987; Brophy and Good 1986; Cohen 1981; Rutter 1983; Weinert, Schrader, and Helmke 1989). Studying how instructional quality impacts SETs, Cashin and Downey (1992, 1999) found that students' overall rating of teacher effectiveness tracked closely with their perceptions of learning, and this opinion finds support in the findings by Ryan and Harrison (1995). Similarly, O'Connell and Dickinson (1993) contend that the correlations between ratings given to instructors and learning was low, but the agreement between ratings and perceived learning was high (also see Clayson 2009).

In their comparison of private and public high schools, Coleman and his colleagues found that academic demands and teaching variables have substantial effects on academic achievement (Coleman, Hoffer, and Kilgore 1982; Hoffer, Coleman, and Greeley 1985). Studies directed towards how such homework and course demands affect SETs cover the full range of possible outcomes. Marsh and Roche (2000) contend that a positive relationship exists between the perception of greater amounts of homework and how students evaluate their teachers. However, the lion's share of research occupies the middle ground in the argument, stating that there is either no, or only a very small correlation between the two (Cohen 1981). Finally, Paswan and Young (2002) position themselves at the other end of the debate by finding that course demands actually have a negative effect on SETs.

The diversity of opinion and inconclusive results outlined in the above, along with a multitude of additional studies, suggests that while the research into this field is overwhelming, it is also far from settled. Neither does it conclusively address students outside North America.

Assessing the specifics of attitudinal variables among Japanese students is

needed to help better position them in the general debate, and provide a more appropriate geographical and cultural base from which SETs can be understood. Such additional work targeted to narrower branches of the discussion will help bring clarity to the forest as a whole, while improving the quality of SETs conducted in Japan in specific. The authors believe that their research is well positioned to assist in doing both.

## **HYPOTHESIS**

The specific addition to existing research addressed in this study is the role of attitudes toward instructional quality and academic demands as predictors of Japanese college students' perceptions of self-improvement in English language courses. We propose that the higher levels of English skills improvement estimated by Japanese college students are explained by their ratings of higher instructional quality and greater course demands. In the analysis, we include respondents' sex, class period, and instructors' sex as controls because other research shows them to be related to language skills improvement (Llach and Gallego 2012 ) and to perceived instructional quality and course demands (Weinert, Schrader, and Helmke 1989).

## **METHODS**

Our objective was to administer questionnaires to a sample of first-year college students in Japan. The questionnaire contained a self-report measure of English language skills improvement, plus measures of perceived instructional quality and outside classwork variables. Control variables, identified from previous research, also were included.

## SAMPLE

Data for this research come from a larger course evaluation survey initiated by one of the authors. In the winter of 2013, Japanese course evaluations were administered to a sample of students in one public university in Japan.<sup>1</sup> The university has a total enrollment (graduate and undergraduate) of approximately 10,000 students and is in an area of about 460 thousand inhabitants that also contains the capital city of the prefecture. The respondents were registered in English I courses taken primarily by first-year students, all of whom had declared a major. Consequently, our respondents were students enrolled in freshman-level English courses across a variety of majors to which the authors could gain access. These included all sixteen majors such as law, economics, chemistry, and medicine representing all three colleges at the university.

A total of 516 questionnaires (written in Japanese) were completed by the students in nineteen English I courses in the university. Six professors had given prior agreement to administer the surveys during class sessions.<sup>2</sup> The questionnaire's cover letter clearly stated that all responses were anonymous and confidential. One respondent who identified himself as a senior was eliminated, resulting in an  $N$  of 515 for the analysis. Among the 515 students in the analysis, 56 percent were male, somewhat fewer than their overall representation of 62 percent at the university.

## MEASURES

### *English Skills Improvement*

Our measure of self-improvement in English language courses is based on students' perceptions of their improvement of English language skills. As Clayson (2009) suggests, such perceived improvement and actual improvement are not synonymous since a student's perception of his/her English skills advancement might not be apparent in actual gains measured objectively in external examinations such as

TOEIC and TOEFL. However, we find rather consistent evidence that student perception of learning is associated positively with student learning (e.g., Baird 1987; Steiner et al. 2006). Therefore, we assess the effects of respondents' perception of the instructional quality and outside classwork on their present estimates of the degree to which their English language skills have improved by taking the course.

To measure perceptions of English skills improvement, respondents were simply asked, "How much do you think your English skills have improved by taking this course?" Responses were given on a four-point scale ranging from "not improved at all" (coded 1) to "improved very much" (coded 4). The variable *English Skills Improvement* has a mean of 3.02 with a standard deviation of .52.

### ***Instructional Quality***

The questionnaire contains three items measuring the quality of instruction: (1) instructional clarity; (2) instructional pacing; and (3) opportunity to participate in student-initiated activities. Instructional clarity is measured with the following question: "Has the instructor explained the material in a way that is easy for you to understand?" The response options were "very uneasy" (coded 1), "somewhat uneasy" (coded 2), "somewhat easy" (coded 3), and "very easy" (coded 4). The variable *Instructional Clarity* has a mean of 3.56, with a standard deviation of .62. For our measure of instructional pacing, respondents were asked the following question: "Has the instructor set aside a reasonable amount of time for you to understand the material?" The response options range from "not reasonable at all" (coded 1) to "very reasonable" (coded 4). The variable *Instructional Pacing* has a mean of 3.56, with a standard deviation of .64. Finally, to measure the opportunity for student-initiated activities, respondents were asked the following question: "How often have you had the opportunity to participate in student-initiated class activities such as pair-work and small group discussion?" The response options were "never" (coded 1), "rarely" (coded 2), "sometimes" (coded 3), and "often" (coded 4). The

variable *Opportunity for Activities* has a mean of 3.59, with a standard deviation of .62.

Our original intent was to create an overall scale of the perceived quality of instruction with all three items. With this goal, these three items were then subjected to principal components and reliability analyses. The reliability analysis, however, indicated that the creation of this single scale of “instructional quality” could not be justified. Instead the *Instructional Quality* scale is the sum of the scores of instructional clarity and instructional pacing with an *alpha* of .70 (*s.d.* = 1.10), leaving *Opportunity for Activities* as a separate variable.

### ***Course Demands***

Our measure of course demands is based on students’ ratings of outside classwork operationalized as the *time*, in hours per week, spent studying outside of class. In reference to the guidelines of the University Establishment Standards (2013), outside classwork is measured by the following three Likert items, each answered on a five-point scale ranging from “no time” (coded 1) to “2 or more hours a week” (coded 5):

How many hours per week have you spent on preparing for the next lessons?

(mean = 2.42, *s.d.* = 1.25).

How many hours per week have you spent on reviewing the previous lessons?

(mean = 1.83, *s.d.* = .83).

How many hours per week have you spent on completing the homework assignments? (mean = 3.18, *s.d.* = 1.04).

A principal components analysis indicated a single factor, with eigenvalues of 1.63, .84, and .53. Cronbach’s *alpha* for the linear composite of the three items is .55. Although the value of .55 is somewhat lower than expected, the three items are sufficiently correlated to produce a general outside classwork scale with an



acceptable reliability ( $\alpha < .50$ ). The variable *Course Demands* has a mean of 7.43, with a standard deviation of 2.29.

### ***Control Variables***

In all analyses, we control for three variables. “School Year” is not one of these since only first-year respondents are included in the analysis. *Respondents’ Sex* is a dummy variable coded 1 for male respondents and 0 for female respondents. The variable *Respondents’ Sex* has a mean of .56 with a standard deviation of .50. Since class period, in terms of the impact of time-of-day instruction on task performance (Wile and Shouppe 2011), might affect respondents’ perceptions of skills improvement, we included a measure of class period as a control variable.<sup>2</sup> Class period is an interval-level variable with response options of “1st period” (coded 1) to “4th period” (coded 4). The variable *Class Period* has a mean of 2.40 with a standard deviation of 1.05. Finally, *Instructors’ Sex* is a dummy variable coded 1 for male instructors and 0 for female instructors. The variable *Instructors’ Sex* has a mean of .41 with a standard deviation of .49.

## **ANALYSAIS AND FINDINGS**

The analysis is performed in two stages. Table 1 assesses all bivariate relationships among the variables we use as predictors. Table 2 then estimates the direct effects of the instructional quality, opportunity for activities, and course demands on English skills improvement. Because direction is predicted in the hypothesis, one-tailed tests are appropriate, and the conventional .05 level is used for judgments concerning significance.

## **CORRELATIONS**

Parts of the bivariate correlation matrices upon which the regressions are based are

reported in Table 1. As expected, English skills improvement is substantially and significantly ( $p < .001$ ) related to all three independent variables. The correlation is especially strong between instructional quality and English skills improvement ( $r = .413$ ). The English skills improvement is also positively and significantly related to both opportunity for activities ( $r = .227$ ) and course demands ( $r = .306$ ). Though not reported in the table, instructors' sex is positively and significantly correlated with English skills improvement ( $r = .142$ ,  $p = .001$ ), indicating that respondents taught by male instructors tend to estimate higher levels of self-improvement of English language skills than those taught by female instructors.

**Table 1. Bivariate Correlations for All Variables ( $N = 515$ ; one-tailed significance tests)**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) English Skills Improvement	-----	-.016 (.361)	.022 (.308)	.142 (.001)	.413 ( $<.001$ )	.227 ( $<.001$ )	.306 ( $<.001$ )
(2) Respondents' Sex (1 = male, 0 = female)		-----	-.048 (.136)	-.024 (.290)	-.063 (.075)	-.105 (.009)	.041 (.177)
(3) Class Period			-----	-.314 ( $<.001$ )	-.168 ( $<.001$ )	-.148 ( $<.001$ )	.210 ( $<.001$ )
(4) Instructors' Sex (1 = male, 0 = female)				-----	.199 ( $<.001$ )	.016 (.363)	-.029 (.256)
(5) Instructional Quality					-----	.305 ( $<.001$ )	-.013 (.385)
(6) Opportunity for Activities						-----	.038 (.198)
(7) Course Demands							-----

## OLS REGRESSIONS

Isolation of independent contributions of the three independent variables on English skills improvement is potentially problematic since there exists a strong correlation between instructional quality and opportunity for activities. Table 1 reveals that the

correlation is positive (.305) and significant beyond the .001 level. However, inspection of multicollinearity diagnostics from SPSS outputs in the various regressions indicated no harmful multicollinearity problem. Thus, the standardized partial regression coefficients (*Betas*) from the OLS regressions for English skills improvement regressed on the three independent variables and the three control variables are reported in Table 2.

**Table 2. Effects of Instructional Quality, Opportunity for Activities, and Course Demands on English Skills Improvement ( $N = 515$ ; one-tailed significance tests)**

<b>Variable</b>	<b><i>b</i></b>	<b><i>Beta</i></b>	<b><i>p</i></b>
Respondents' Sex (1 = male, 0 = female)	.014	.014	.361
Class Period	.035	.071	.043
Instructors' Sex (1 = male, 0 = female)	.102	.096	.009
Instructional Quality	.178	.376	<.001
Opportunity for Activities	.093	.111	.003
Course Demands	.067	.294	<.001
<i>(intercept)</i>	.793		
$R^2$	.285		
<i>p</i>	<.001		

As expected, instructional quality and course demands have significant ( $p < .001$ ) positive effects on English skills improvement, with the effect of instructional quality ( $Beta = .376$ ) being slightly greater than the effect of course demands ( $Beta = .294$ ). Opportunity for activities has a significant positive effect, consistent with the prediction, though the magnitude of the effect is much smaller than the effects of instructional quality and course demands. Thus, consistent with theoretical expectations regarding the determinants of student learning, we find that students' perceptions of English skills improvement correspond to their perceptions of the quality of instruction, opportunity for student-initiated activities, and time spent on studying outside of the class.

Among the control variables, instructors' sex has a significant direct effect on English skills improvement ( $Beta = .096$ ,  $p = .010$ ), and it is positive as would be

expected from the significant positive correlation between the two in Table 1. The positive sign indicates that respondents taught by male instructors tend to perceive higher levels of English skills improvement than those taught by female instructors. This is consistent with the findings in Table 1, where respondents taught by male instructors score higher on English skills improvement. Furthermore, a surprising finding in the equation is the significant *positive* direct effect of class period on English skills improvement, indicating that those who are taking courses in later class periods tend to estimate higher levels of English skills improvement. Recall that the bivariate correlation between these two variables ( $r = .002$ ) in Table 1 is insignificant. But with controls for the other variables, class period has a significant positive direct effect ( $Beta = .071, p < .05$ ), and a more detailed analysis (not reported here) reveals that it is much greater ( $Beta = .141, p = .001$ ) with the combination of instructional quality and opportunity for activities. In other words, once instructional quality and opportunity for activities are controlled, courses offered in later periods significantly *increase* the levels of English skills improvement. Apparently the significant inverse correlations of class period with instructional quality ( $r = -.168$ ) and opportunity for activities ( $r = -.148$ ), plus the positive effects of these two independent variables on English skills improvement mask the unexpected significant positive effect of class period on English skills improvement.

In summary, the analysis suggests that Japanese college students' perceptions of instructional quality, opportunities for activities, and course demands contribute to their self-assessment of English language skills improvement. These effects remain significant even with controls for class period and for respondents' and instructors' sexes as possible sources of spuriousness. Class period, contrary to our initial speculation, actually has a positive effect on English skills improvement, but only when the other variables are controlled.

## DISCUSSION

The fact that students' ratings of the instruction and higher work demands of the course have positive effects on their perceptions of their overall English skill is significant in the aggregate, but finds particular relevance across the Japanese university system.

Instructors may find it useful to consider the research from not only an academic perspective, but also from the standpoint of their individual career situations. SETs are increasingly used as a component of teacher reviews, and their outcomes have the potential to affect salary, promotion, and therefore, the instructor's career as a whole. The role of SETs in job evaluations is especially pronounced in Japan, where a substantial portion of English courses are taught by either non-tenured or part-time instructors. As these jobs are not tenured and tend to be dependent on performance, SET results take on an even greater importance. The O'Connell and Dickinson study (1993) found that students' perceptions of what they had learned had a much higher correlation with how they ranked classes than did the differentials in their pre- and post-test scores. Therefore, it is entirely possible that an instructor's evaluation could be more dependent upon how the students perceive the class than on what they actually learn. With this in mind, instructors may find advantage in considering these findings as they design English curricula.

Individual departments and even the university as a whole may also find value in considering these results when setting English education guidelines. SET outcomes reflect not only on the individual instructors, but on the department and the entire institution. Through targeted adjustments to curricula standards mindful of future SET feedback, overall levels of course satisfaction can be modified upward.

Lastly, and most importantly, the study provides data useful for improving Japanese students' performance in English education. With regard to SETs, it has been shown that perception often is reality. Not only is there a correspondingly high level of *perceived* learning when students give high ratings for instruction, but

examinations define correspondingly high levels of *real* learning (Cohen, 1981; Feldman, 1989; Marsh, 1987). Whether the learning is perceived or actual, higher student evaluations can play a role in improving student confidence and providing motivation to engage in further studies. Consequently, active consideration of the findings that Japanese college students' perceptions of instructional quality, opportunities for activities, and course demands affect self-assessment of their English language skills improvement can have significantly positive implications for instructors, university administration, and English language students.

### **ACKNOWLEDGMENTS**

We are thankful to the instructors who participated in collecting the data for their invaluable contribution to this research.

### **NOTES**

- <sup>1</sup> First-year college students were chosen as respondents for two reasons. First, we had easier access to them in the university than to students in upper-division courses. Second, first-year students are required to take English I courses to meet one of the general education requirement at the university. We realize, of course, that students in upper-division courses are excluded from our research design and might perceive higher (or lower) levels of English skills improvement than first-year students at the university.
- <sup>2</sup> Each class meets for 90 minutes with a 15-minute break between classes except for a 60-minute lunch break (12:00-13:00) between 2nd and 3rd periods.

### **REFERENCES**

- Amaki, Y. (2008). Perspectives on English education in the Japanese public school system: The views of foreign assistant language teachers (ALTs). *Education Studies in Japan: International Yearbook*, 3, 53-63.
- Baird, J. S. (1987). Perceive learning in relation to student evaluation of university

- instruction. *Journal of Educational Psychology*, 79, 90-91.
- Brandenburg, G. C., & Remmers, H. H. (1927). A rating scale for instructors. *Educational Administration & Supervision*, 13, 399-406.
- Brandenburg, D. C., Slinde, J. A., & Bastista, E. E. (1977). Student ratings of instruction: Validity and normative interpretations. *Research in Higher Education*, 7, 67-78.
- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and industrial performances*. San Francisco, CA: Jossey-Bass.
- Brophy, J. E. (1986/1987). Teacher effects research and teacher quality. *Journal of Classroom Interaction*, 22, 14-23.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (Vol. 3, pp. 328-375). New York: Macmillan.
- Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall and J. Franklin (Eds.), *Student ratings of instruction: Issues for improving practice*. New directions for teaching and learning, no. 43. San Francisco, CA: Jossey-Bass.
- Cashin, W. E., & Downey, R. G. (1992). Using global student rating items for summative evaluation. *Journal of Educational Psychology*, 84, 563-572.
- Cashin, W. E., & Downey, R. G. (1999, April). Using global student rating items for summative evaluation: Convergence with a second criterion. Paper presented at *the Annual Meeting of the American Educational Research Association*, Montreal, Quebec, Canada.
- Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco, CA: Jossey-Bass.
- Clayson, Dennis E. (2009). Student evaluations of teaching: Are they related to what students learn?: A meta-analysis and review of the literature. *Journal of Marketing Education*, 31, 16-30.
- Cohen, P. A. (1981). Student-ratings of instruction and student-achievement - a

- meta-analysis of multisection validity studies. *Review of Educational Research*, 51, 281-309.
- Coleman, J. S., Hoffer, T., & Kilgore, S. (1982). *High school achievement: Public, Catholic, and private schools compared*. New York: Basic Books.
- Costin, F., Greenough, W. T., & Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, 41, 511-536.
- Feldman, K. A. (1978). Course characteristics and college students' ratings of their own teachers: What we know and what we don't. *Research in Higher Education*, 6, 223-274.
- Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education*, 30, 583-645.
- Greimel-Fuhrmann, B., & Geyer, A. (2003). Students evaluation of teachers and instructional quality: Analysis of relevant factors based on empirical evaluation research. *Assessment and Evaluation in Higher Education*, 28, 229-238.
- Hoffer, T., Coleman, J., & Greeley, A. (1985). Achievement growth in public and Catholic high school. *Sociology of Education*, 58, 74-97.
- Llach, M. P. A., & Gallego, M. T. (2012). Vocabulary knowledge development and gender differences in a second language. *Estudios de Linguística Inglesa Aplicada*, 12, 45-75.
- Marsh, F. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H. W., & Roche, L. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92, 202-228.



- McFadden, K., & Dart, J. (1992). Time management skills of undergraduate business students. *Journal of Educational Psychology, 82*, 760-768.
- Murray, H. G. (2005, June). Student evaluation of teaching: Has it made a difference? Paper presented at *the Annual Meeting of the Society for Teaching and Learning in Higher Education*, Charlottetown, Prince Edward Island, Canada.
- Nois, S. A., & Hudson, G. I. (2006). Academic performance of college students: Influence of time spent studying and working. *Journal of Education for Business, 81*, 151-159.
- O'Connell, D. Q., & Dickinson, D. J. (1993). Student ratings of instruction as a function of testing conditions and perceptions of amount learned. *Journal of Research and Development in Education, 27*, 18-23.
- Ory, J. C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions for Institutional Research, 109*, 27-44.
- Paswan, A. K., & Young, J. A. (2002). Student evaluation of instructor: A nomological investigation using structural equation modeling. *Journal of Marketing Education, 24*, 193-202.
- Remmers, H. H. (1927). The Purdue rating scale for instructors. *Educational Administration and Supervision, 6*, 399-406.
- Remmers, H. H. (1928). The relationship between students' marks & students' attitudes toward instructors. *School & Society, 28*, 759-760.
- Remmers, H. H. (1930) To what extent do grades influence student ratings of instructors? *Journal of Educational Psychology, 21*, 314-316.
- Remmers, H. H., & Brandenburg, G. C. (1927). Experimental data on the Purdue ratings scale for instructors. *Educational Administration & Supervision, 13*, 519-27.
- Rodin, M., & Rodin, B. (1972). Student evaluations of teachers. *Science, 177*, 1164-1166.
- Rutter, M. (1983). School effects on pupil progress: Research findings and policy

implications. *Child Development*, 54, 1-29.

- Ryan, J., & Harrison, P. (1995). The relationship between individual instructional characteristics and the overall assessment of teaching effectiveness across different instructional contexts. *Research in Higher Education*, 36, 577-594.
- Seldin, P. (1993). How colleges evaluate professors: 1983 versus 1993. *AAHE Bulletin* Oct., 6-8.
- Steiner, S., Holley, L. C., Gerdes, K., & Campbell, H. E. (2006). Evaluating teaching: Listening to students while acknowledging bias. *Journal of Social Work Education*, 42, 355-376.
- University Establishment Standards. (2013). *University Establishment Standards*. Retrieved October 7, 2013 from ([http://law.e-gov.go.jp/cgi-bin/idxselect.cgi?IDX\\_OPT=3&H\\_NAME=&H\\_NAME\\_YOMI=%82%A0&H\\_RYAKU=1&H\\_CTG=1&H\\_YOMI\\_GUN=1&H\\_CTG\\_GUN=1&H\\_NO\\_GENGO=S&H\\_NO\\_YEAR=31&H\\_NO\\_TYPE=5&H\\_FILE\\_NAME=S31F03501000028](http://law.e-gov.go.jp/cgi-bin/idxselect.cgi?IDX_OPT=3&H_NAME=&H_NAME_YOMI=%82%A0&H_RYAKU=1&H_CTG=1&H_YOMI_GUN=1&H_CTG_GUN=1&H_NO_GENGO=S&H_NO_YEAR=31&H_NO_TYPE=5&H_FILE_NAME=S31F03501000028)).
- Weinert, F. E., Schrader, Friedrich-W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, 13, 895-914.
- Wile, A. J., & Shoupe, G. A. (2011). Does time-of-day of instruction impact class achievement? *Perspectives in Learning: A Journal of the College of Education & Health Professions*, 12, 21-25.

# 大学生の英語学習成果の規定要因

— 大学1年生対象の授業評価アンケート調査結果より —

小林 恵美子  
ビントリフ ジョン

## 要旨

大学教育の質を保証するためには、データに基づき、客観的に学生の学習成果を把握することが重要である。これまで米国では、学生の学習成果に関する研究が盛んに行われ、その成果を踏まえた学生調査が組織的に実施されている。しかしその一方で、日本における研究は数少ない。多くの大学が学生の学習目標を定め始めており、実際にどれだけの学習成果が上がったのかを測る取り組みも進めてはいるが、学生の学習成果向上の仕組みについては、いまだ十分解明されてはいない。そこで本稿では、米国の学習成果に関する研究で用いられている枠組みを踏まえ、大学生の英語学習成果の規定要因について検討する。具体的には、大学生の英語運用能力における学習成果に着目し、授業の進め方、学習への取り組み方が学生の英語運用能力向上感にどのような影響を与えているのかについて、授業評価アンケート調査結果を用いて実証的に解明する。国立大学1年生対象の19クラス515名からのデータを重回帰分析したところ、担当教員の授業の進め方が効果的であるほど（説明の明快さ、授業進行速度の適切性、学生参加機会の多寡）、そして授業外の学習時間が長いほど、学生は自身の英語運用能力の向上を感じていることが確認できた。