University accreditation in Japan: Problems and possibilities for reforming EFL education

Bern Mulvey
Iwate University

As of 2004, all universities in Japan must submit to an external accreditation evaluation, to be repeated every seven years. The universities are to receive detailed written assessments in multiple categories from one of four official accrediting agencies—the Japan University Accreditation Association (JUAA), the National Institution for Academic Degrees and University Evaluation (NIAD-UE), the Japan Institution for Higher Education Evaluation (JIHEE), and the Japan Association for College Accreditation (JACA). These assessments are to be publicized. The universities also receive grades: pass, probation, and fail.

Japan’s Ministry of Education (hereafter MEXT) has repeatedly (e.g., Chuou Kyouiku, 2002: Daigaku Shingikai 1991, 1998; MEXT 1999, 2001, 2003, 2004a, 2004b; Rinji, 1986, etc.) made its intentions clear: Through these new requirements, it hopes to induce systemic improvements in teaching and research quality, not to mention encourage technical upgrades and on-campus diversity. Universities are being prodded into a greater level of transparency, not just in regards to finances and accounting, but in their grading and advising policies and even hiring practices. Safeguards (e.g., procedures to prevent or punish various types of harassment) for students, staff, and faculty have to be adopted—or clarified and strengthened in the case of universities with policies predating 2004. Finally, these accreditation assessments are to serve the additional function of confirming both the presence of on-campus Faculty Development (hereafter FD)
committees and the effectiveness of their various activities.

In other words, at least in theory, university accreditation represents an unparalleled opportunity to achieve meaningful, lasting educational reform in this country. However, as is often the case with reform attempts of this scope, the reality is much more complex and, particularly with regards to EFL classes and their (often non Japanese) instructors, troubling.

This paper provides a critical overview of the accreditation requirements, their impact, and the potential opportunities they represent. Having served as Dean at a Japanese university undergoing accreditation, headed a university accreditation committee, and authored major sections of the 100-page official accreditation report, my discussion includes first-hand observations not elsewhere available in English. I hope that my commentary here will lead to a more informed public dialogue about Japanese-style university accreditation—both its problems and possibilities.

**Background**

University accreditation became official policy in Japan with the passing of the following amendments (MEXT, 2004a, 2004b, 2004c): Article 69.4.1-2 of the Gakkou Kyouikuhou makes accreditation a legal requirement, not to mention codifies the official standards and assessment procedures; Article 69.3.2 requires that all 4-year universities in Japan submit themselves to external evaluation every seven years; Article 69.3.4 requires that all evaluation results be publicized. The reach of these amendments, particularly for the former national universities, is reinforced by Article 16.1 of the Dokuritsu Gyousei Houjin Daigaku Hyouka/Gakui Juyo Kikouhou, which states in part that all universities must now undergo periodic assessments of education and research quality, with these external evaluations paid for by the universities themselves and the results publicized afterward (MEXT, 2004b, 2004c).

These new requirements emanated partially out of MEXT’s desires both to ensure a uniform level of quality throughout its higher education system and to increase Japan’s educational competitiveness worldwide (see Chuo Kyouiku 2004; Daigaku Shingikai 1991, 1998; MEXT 1999, 2001; Rinji, 1986, etc.). Note that limited, informal university accreditation had existed in Japan from 1947, when 47 universities banded together to create JUAA (JUAA, 2009)—though, as fitting the founding principle of the latter (“We will use our members’ independent efforts and mutual cooperation among them to measure improvements in university quality”), evaluation was conducted for and by members (and later would-be members) according to standards they themselves set, with (until the 1995 publication of their accreditation manual) almost no transparency outside the organization. In April of 1986, the Rinji Kyouiku Shingikai, citing a need “to rebuild trust in the education world” (教育界の信頼の回復, p. 16), first advocated the creation of a government-sponsored, standardized evaluation process (p. 88), both in the form of required “self assessments” (自己評価) and formal “accreditation”—the latter term originally Katakana-ized as アクレディテーション (instead of the now standard “認証評価”), underlining the borrowed nature of the concept (see also Chuo Kyouiku, 2002).

Self-assessments were begun immediately: 88% of universities had conducted at least one by 1997, with 65% publicizing the results (Daigaku Shingikai, 1998). However, the quality of these self-evaluations was heavily criticized; e.g., they were described by MEXT (Daigaku Shingikai, 1998) as being mere “inspections without any assessment” (点検あって評価なし), with universities seen as unwilling and/or unable to achieve the levels of transparency and self-honesty sufficient to allow for productive self-assessment. Requiring evaluations by government-approved external agencies was meant to change this. The various assessment categories and the overall accreditation process are closely based on accreditation in the U.S. (Chuo Kyouiku, 2002; Rinji, 1991, etc.), and with a similar intent. In other words, by requiring “objective” (客観性) external evaluations ostensibly necessitating complete institutional “transparency” (透明性), MEXT hoped to force self-improvement on the weaker programs and stimulate a general raising of standards among all educational institutions.

However, it should be kept in mind that MEXT is also motivated by potentially devastating de-
mographic and economic pressures. Between 1950 and 1980, Japan's population increased 40%, from 83 million to 117 million (Abe, 2000, p. 47). Rapid economic growth experienced over the same period, combined with policy changes enabling easier university creation (especially private), led to an explosive expansion of new institutions. In 1949, the number of universities and students stood at 178 universities and 130,000 respectively; this increased to 382 and 1.4 million by 1970, and 507 and 2.2 million in 1990 (Abe, 2000, p. 48). More importantly, despite the decreasing birthrate (and, after 1990, a corresponding decrease in overall student numbers), the number of Japanese universities has actually continued to increase: up to 756 in 2007, with approximately 2.8 million students enrolled (MEXT 2007a, p. 5).


Japan's traditional university feeder programs have reached the point where they can no longer graduate a number of students sufficient to maintain the economic vitality of the majority of Japan's universities. In 1999, for example, 756,149 high school students applied for entrance into university, with 602,078 accepted. This works out to a success rate of 80%—an all-time high, and one which is expected to climb even higher over the next few years. (p. 13)

This success rate climbed to 90.5% in 2006 (MEXT, 2007). That same year, 47.1% of Japan’s private universities (266 schools total) reported that they failed to meet their enrollment goals (Nihon Shiritsu, 2006, p. 23). Moreover, Daigaku Shingikai (2000, p. 2), MEXT (1997, 1998), and Mulvey (2001, pp. 13) have estimated that the applicant-to-university space ratio (志願者に対する収容力) will reach exactly 1:1 by the year 2009.4 In other words, by as early as this year, there will be exactly one student applicant for each available university (including 2-year college) berth—meaning that, particularly in the case of two-year and many regional 4-year institutions, rejecting any student admission applications will have a direct economic impact (see also MEXT, 2007, and Nihon Shiritsu, 2006).

Consequences predicted by the Japanese government include faculty layoffs, school closures, hiring freezes, and severe budget reductions (Jannuzi & Mulvey, 2000; Mori, 1999; Mulvey, 2000, 2001). Moreover, as the number of applicants comes to equal the number of places available, post-secondary programs have been forced to relax admissions standards in order to maintain enrollment levels sufficient to ensure their economic viability. The Chuuo Kyouiku Shingikai (“Misu-machi,” 1999), Kariya (2002), MEXT (2000c, 2008) and the Nihon Keizai Shinbun (“Daigaku de Fueru,” 1999), for example, all describe a resultant decline in median academic ability vis-à-vis test-related skill areas among even successful university entrants, with many of the freshmen evaluated lacking minimal skills in math, English, the sciences, and the Japanese language. Daigaku Shingikai (2000, pp. 1-4) also found that a growing number of universities are being forced to accept students even with extremely low examination results (see also “Cram,” 1998; Mulvey 2001).

With respect specifically to English education, another side effect of this process has been the weakening (if not elimination) of the potential for beneficial exam washback-type influence on high school teaching methodology and textbook content. The long history of problems with the various types of entrance exams is well documented. However, since at least 1987 (Daigaku Shingikai, 2000), MEXT has been advocating revisions to exam content and overall selection procedures. Indeed, Guest (2000, 2008) and Mulvey (1998, 1999, 2001) cite over seventy, mostly Japanese, studies which underline the changes/improvements made to the various entrance examinations—e.g., the widespread inclusion of writing and listening sections, not to mention a de-emphasis on translation or grammar-focused discrete-item problems—yet with no evidence of a corresponding, systemic change to the high school English curriculum in this country. Mulvey (2001) concludes:

As even students with extremely poor exam results can now successfully gain admittance into many universities, the pressing need for high school educators to make curriculum
adjustments with every change in exam content would seem to be removed. This in turn would make high school curriculums less open to exam influence, as many students would still pass regardless of any preparation deficiency in their high school English classes. In other words, perhaps the lack of exam influence described earlier is the result of the examinations evolving where high school curriculums have not? (p.14)

The Chuo Kyouiku Shingikai (1999) makes similar observations, concluding, “A number of high school educators continue to hold the opinion that, until the entrance examinations change, the curriculum cannot be changed. Well, we would like them to understand that, both with the Center exam and the individual university entrance exams, extensive reforms have been ongoing for some time.”

Keep in mind that holding back academically struggling students, let alone expulsion for failing grades, is legally impossible through the end of junior high school (when compulsory education finishes in Japan); on average, 97.7% of these students choose to enter high school (MEXT, 2007a, 2007b), with graduation rates extremely high (e.g., in 2005, 97.9% of the high school population graduated on time, with problems with school life/peers the main reason [38.6%] given by the few students who did not—see MEXT, 2006.) Accordingly, in many ways, the entrance exams (both to universities and to the more academically oriented high schools) have long served as an important source of objective evaluation and selection in Japan, ensuring—however inadequately—a level of quality control throughout the system. As the number of test applicants approaches the breakeven point necessary to ensure the continuing financial viability of the various schools, this main function will increasingly weaken in effectiveness and need to be replaced.

The accreditation process

Here, I will focus on the process and specific requirements for JIHEE (2009a, 2009b), though the policies and procedures at all four agencies are necessarily similar, the general framework being mandated by law (see MEXT, 2004a, 2004b).

One year before they wish to be evaluated, universities each must submit an application form and fee to the accreditation agency. After that, universities begin work on the centerpiece of the accreditation process: the 100-page self-assessment report submitted by each university in the summer preceding the three-day onsite inspection. This report requires an extremely detailed and well-documented self-review, with extensive Self-Evaluation (自己評価) and concrete Measures/Future Plans for Improving and Enhancing Quality (改善・向上方策「将来計画」) needed in response to each standard. The report is submitted by the end of June; at about this same time, four to six reviewers (all professors and/or senior administrators) are assigned to the university to conduct the onsite inspection. These reviewers each receive a copy of the self-assessment report, to which they respond formally with written follow-up questions (書面質問) by the end of August.

While minor differences among the four accrediting agencies exist, the substance of the assessment categories remain unchanged from those first proposed by MEXT for self-evaluation (Chou Kyouiku 2006a, Daigaku Shingikai, 1991, pp. 36-38). The following are from the JIHEE website:

The sample questions referred to above are actual follow-up questions asked by accreditation referees in response to university-submitted Self-Assessment Reports. The questions underline some of the key referee concerns: the importance of teamwork, fairness, safety, student-focused teaching and research, faculty voice (including participation in the administrative process), and the necessity for complete transparency and honest reflection at all levels. At my university as well, we received a similar line of questions, including:

- Your FD Committee, in charge of pedagogy and the practical application educational tools or materials, has conducted how many seminars, with how many faculty participating at each?
- With regards to your EFL classes, we note that you leave class content and methodology to the discretion of each individual faculty member. Are there any plans to systematize and unify your efforts?
## Table 1. JIHEE accreditation standards

<table>
<thead>
<tr>
<th>Standard</th>
<th>Sample follow-up questions in each category</th>
</tr>
</thead>
</table>
| 1. Philosophy behind Establishment, Mission, and Objectives of the Institution | 1. What steps have you taken to publicize your university’s mission?  
2. What role did faculty and staff play in developing or articulating the university’s mission?                                                                                                                                     |
| 2. Education and Research Organization                                  | 1. How are the General Education courses organized?  
2. In addition to teaching, what role(s) do faculty play at your institution?                                                                                                      |
| 3. Curriculum                                                           | 1. Are syllabi on file for each course?  
2. Are class activities and contents accurately reflected by the syllabi?                                                                                                             |
| 4. Students                                                             | 1. How are student evaluation results being used to improve academic quality?  
2. What has been the pattern of status change [e.g., withdrawal and/or leave of absence] with regards to enrolled students, and what is being done to correct any problems? |
| 5. Faculty                                                              | 1. With regards to hiring and promotion, how are teaching accomplishments evaluated?  
2. Are Faculty Development activities coordinated and systemic?                                                                                                                  |
| 6. Staff                                                                | 1. How is “on the job training” handled?  
2. How is worker performance evaluated and what is done with the evaluations?                                                                                                     |
| 7. Administration                                                        | 1. Is there a system in place to insure that the results from the various self-assessments are practically applied to improving the school?                                                                                                  |
| 8. Finance                                                              | 1. Could we have some more information about current and future plans to publicize the university’s financial records?                                                                                                                    |
| 9. Educational and Research Environment                                 | 1. What has been done to make the university “barrier free”?  
2. Has any preparation been made with regards to potential natural emergencies?                                                                                                           |
| 10. Societal Interaction                                                 | 1. What is the level of autonomous faculty and staff contribution—for instance through lectures and/or committee membership—in the community and nationally?                                                                                   |
| 11. Social Accountability                                                | 1. What precautions are in place to prevent sexual harassment?  
2. What procedures are in place in case of a sexual harassment complaint?                                                                                                      |
Does offering a 2-year renewable contract impact faculty performance and feelings of belonging at the university?

Again, the emphases were on our educational objectives and methodology, not to mention our treatment of both students and faculty. How did we hope to foster critical thinking, develop problem-solving strategies, and/or prepare students to function effectively in an information economy in an internationalizing world? How did our curriculum help our students achieve these objectives? Our treatment of especially foreign faculty, including issues of workload and pension or retirement benefits, was an important topic as well.

Indeed, throughout the 2-year process, I was struck repeatedly by the potential benefits, the possibilities for true reform, that accreditation represents. Keep in mind that, at least in theory, there is more at stake for these universities than critical written assessments and/or bad publicity. The external auditors also assign grades: pass (認定—literally “recognized”), probation (保留) and fail (不認定—literally “not recognized”). Assuming criteria appropriateness, objective assessment and enforcement, universities can now be held accountable for what they teach and how, not to mention their handling of student and teacher evaluations, hiring, and firing.

Accreditation questions…and opportunities

While impressive overall, there remain areas of concern, particularly related to English language teaching and foreign faculty. Much of the problem lies in the specifics of MEXT’s chosen methods to improve and evaluate, both teachers and teaching. As mentioned earlier (Daigaku 1998, also see MEXT 2006b, 2008), FD is the centerpiece of MEXT efforts to improve quality at the university-level, with the accreditation reviews intended both to document the incorporation of the FD committees and to evaluate their productivity. MEXT (Chuo Kyouiku, 2006b) defines FD thus:

FD is a generic term for organized, institution-wide efforts to improve class content and teaching quality. This definition is necessarily extremely vague; specific examples, however, include requiring peer-review of teaching, holding faculty study sessions to discuss teaching methodology, and offering orientation seminars for new faculty.

In other words, FD committees in Japan are expected to concentrate solely on improving teaching, though exactly what these improvements should constitute—or even an agreed upon definition of “good” teaching—have yet to be articulated. Moreover, FD effectiveness at each institution is graded during the accreditation reviews according to numerical values—e.g., the total number of activities conducted, not to mention the various percentages of participation (the percent of faculty who have a syllabus for each class, the percent who attend each FD activity, the percent whose classes are evaluated by students, etc). This has, understandably, led to an emphasis on “quantity” over “quality.” For instance, Miyazaki International College, despite its small size and the lack of demonstrable need, had to increase dramatically its number of FD activities; the resulting sense of unnecessary burden was heightened by the redundant and/or overlapping responsibilities of other, similar, committees predating the advent of FD—in addition to the MEXT-mandated FD Committee, thirty-three faculty had to staff, attend meetings, and plan activities for the Committee on Faculty, Committee on Curriculum, Faculty Review Committee, Teacher Education Committee, and Committee on Students/Admissions (among others). Consolidating and/or otherwise focusing/reducing the total number of committees is not an option; all committees are “counted,” all have to demonstrate to the accreditation reviewers a separate, active, agenda.

This emphasis on the quantifiable extends to a heavy reliance on student course evaluation numbers as well. While research in the U.S. (e.g., Felder and Brent, 2004) has demonstrated “a high level of validity” with student evaluations in general (p. 200), several Japan-specific practices are worrisome. Ryan (1998) contains several studies which seem to demonstrate that Japanese students use different internal criteria when rating foreign and Japanese faculty, with the former “not seen as serious teachers” (p. 11). Moreover, typically a single committee, composed of faculty representatives from a cross spectrum of the various departments
(yet unified by a uniform lack of formal training in evaluation method and adult-level teacher education), is put in charge of creating, administering, and assessing student evaluations. Can student-centered, active learning, or both approaches be evaluated accurately and fairly by the resulting forms? As Centra (1993) comments also, the usual result of such collaboration is a student rating form “devised to reflect effectiveness in lectures, lecture and discussion, and other teacher-centered methods” (pp. 47-8).

This is particularly relevant to EFL, particularly the General Education (教養・共通教育) classes that often are the sole prerogative of foreign faculty. These classes tend to be larger than average, dominated by non-majors in the subject area, and among the few (sometimes only) classes on campus taught in the students’ L2. Accordingly, the methodology, materials, and educational goals for these classes necessarily differ from most other courses—even from upper-division English seminar classes (the latter usually smaller, taught mainly in Japanese by Japanese faculty to English majors). Theall (2005, p. 2) notes that “teachers assigned a predominantly heavy load of lower-level, introductory, required, large courses” tended to have “depressed ratings.” Moreover, at most universities, the overwhelming majority of foreign EFL faculty are either part-time hijoukin or on term-limited contracts—Winskowski (2005, p. 41) cites a number of studies which suggest that regular faculty tend to receive higher average evaluations than non regular faculty. Yasuoka (2006), in a MEXT-sponsored study, reinforces these findings. Seminar classes here have the highest average scores (4.3/5), while General Education classes have the lowest evaluation scores (3.8/5). Students with low English scores tend to be (slightly) harsher when evaluating their English teachers. Finally, instructor age significantly impacts evaluation results in Japan: Faculty receive average scores of 3.92 at the age of 30, 3.80 at the age of 40, 3.70 at the age of 50, and 3.58 at the age of 60. Indeed, it can be argued that the current evaluation system (including question content) is skewed against older teachers attempting to utilize communicative, active learning, or both techniques in the larger General Education classes—an unfortunate reality for foreign EFL faculty looking to make a career here.

This is important, as Amano and Nanbu (2005, p. 234) and Ryan (1998, p. 9) note also, because of the aforementioned, accreditation-driven tendency towards increased reliance on evaluation score results in faculty promotion, hiring and retention—and even research funding—decisions. Faculty with lower than average evaluation scores will necessarily—even understandably—be penalized, though this but underlines the lack of input most foreign EFL faculty have into the policy decisions which shape their lives.

Sadly, these are not merely abstract issues. In 2007, Akita Kokusai Kyouyou Daigaku suddenly dismissed almost 20% of its total faculty—all but one non-Japanese with (at least) MA degrees, substantial publications and extensive teaching experience. While I do not presume to know the specific details behind each individual termination, I do know that the teachers collectively appealed this decision to the local labor board and won a preliminary judgment against the university (“Kokusai,” 2007). Questions regarding the treatment of foreign teachers are certainly not confined to Akita; Hall (1998) documents a pattern of concerns dating back over 100 years.

Surely the accreditation reviews, ostensibly designed to ensure curriculum and evaluation improvements, not to mention promote transparency, fairness, internationalization, and modernization, can also play a part in alleviating this situation? Indeed, some evidence exists to suggest that the accreditation assessments are already having a positive impact. To name just two examples, Miyazaki Kokusai Daigaku has used feedback from its assessment review to begin investigating clearer delineations of level- and learner-appropriate learning outcomes for its EFL classes, not to mention better coordination between these classes and the rest of its curriculum. Furthermore, influenced at least partially by its own accreditation review, Iwate Daigaku has chosen to do away with race-based discriminatory hiring. Effective this year, faculty hired full-time will be treated the same regardless of nationality.

These two examples are offered here not to prove the efficacy of the accreditation process, but to demonstrate a cause for hope in that process. Certainly, serious questions remain about the nature and accuracy of the measurement tools.
and the standards of quality being applied—especially with regard to EFL classes and language faculty. Still, in the demands of the accreditation assessments, particularly the requirements of transparency and accountability, lie perhaps our best chance for future, systemic improvements in this area as well.

References


(Endnotes)

1. Trial evaluations were conducted from 2001.
2. JACA deals primarily with 2-year colleges.
3. 会員の自主的的努力と相互的援助によってわが国における大学の質的向上をはかる。
4. I prefer using applicant-to-university space numbers as opposed to other types of data because Japanese students on average apply to multiple (3+) universities, artificially inflating the overall figures. In other words, the former data tells us the real number of applicants who applied to at least one university—and not the total number of applications they sent. From experience as a university administrator, the initial excitement generated from receiving more applications than you have seats fades quickly come March when as many as 40% inform you that they will go elsewhere.
5. 高等学校関係者の中には、「大学入試が変われば高校教育は変わらない」という意見もあるが、大学全体としても、個々の大学においても銳意改革が進められている状況を十分理解してほしい。
6. The questions are taken from 大学評価の心, a JIHEE presentation by Kubo Takeshi on 30 June, 2006.
7. 教員が授業内容・方法を改善し向上させるための組織的な取組の総称。その意味するところは極めて広範にわたるが、具体的な例としては、教員相互の授業参観の実施、授業方法についての研究会の開催、新任教員のための研修会の開催などを挙げることができる。
8. Regarding the latter issues, Amano and Nanbu found the number of universities using evaluations in such a manner “surprisingly large” (むしろ意外に多いとも思える結果であった), though five years later, the number of such schools continues to increase.
9. AIU refused to pay, continuing to claim they had done nothing illegal. As the plaintiffs had by that point all found new employment with better (including tenured/tenure-track) conditions, the case—unfortunately, in my opinion—never went to formal trial.

Bern Mulvey has recent articles in the American Language Review, the Times Higher Education Supplement and Asahi Shimbun, with another forthcoming in the Japan Studies Review. The Fat Sheep Everyone Wants, his book of poetry, won the 2007 Cleveland State University Poetry Center Prize and was nominated for the Pulitzer Prize. Formerly the Dean of Faculty at Miyazaki Kokusai Daigaku, he currently is an associate professor at Iwate Daigaku.

Is your membership due for renewal?
Check the label on the envelope this TLT came in for your renewal date, then go to <jalt.org/main/membership> and follow the easy instructions to renew. Help us to help you! Renew early!