

Issues, Trends and Recommended Practice in  
College Foreign Language Outcomes Assessment: An Interview with  
John McE. Davis, University of Hawai‘i at Mānoa

***PLJ- Please talk about your dissertation research. Why did you decide to write about evaluation/assessment in FL education?***

I have been interested in foreign language (FL) program evaluation and assessment<sup>1</sup> throughout my Ph.D. studies, though the specific idea for the dissertation came via my association with the Foreign Language Program Evaluation Project (FLPEP) at the University of Hawai‘i at Mānoa. FLPEP is a federally funded research group formed to help college foreign language educators engage in evaluation and assessment. FLPEP provides various types of evaluation and assessment-related support to FL educators. Personnel associated with the group deliver workshops and presentations, undertake FL assessment/evaluation research, generate and compile assessment/evaluation publications, and so on. Importantly, much of this activity—and much of my own assessment outreach and consultancy work—has been particularly focused on helping foreign language educators meet increasing demands for student learning outcomes assessment, a now prevalent phenomenon affecting all of U.S. post-secondary education.

Such requirements will be well known to most FL educators in that they have probably experienced requests from their institution (e.g., from a dean or provost) to start outcomes

---

<sup>1</sup> I use the terms ‘evaluation’ and ‘assessment’ following Norris’ (2006) distinctions. According to Norris, evaluation is “the gathering of information about any of the variety of elements that constitute educational programs, for a variety of purposes that primarily include understanding, demonstrating, improving, and judging program value. Evaluation brings evidence to bear on the problems of programs...” (p. 579). Assessment—differing slightly in meaning—is “the systematic gathering of information about student learning in support of teaching and learning. It may be direct or indirect, objective or subjective, formal or informal, standardized or idiosyncratic ... [and] ... provides locally useful information on learners and on learning to those individuals responsible for doing something about it” (Norris, 2006, p. 579). Evaluation in Norris’ taxonomy, then, is a superordinate category encompassing assessment. That is to say, assessment is a type of evaluation in that it is a similarly empirical, investigatory activity though one aimed specifically at analyzing student learning, whereas evaluation might investigate any type of program-related phenomenon.

assessment in their programs. The impetus behind these requests has to do with outcomes assessment functioning as an accountability-system, which accreditation agencies use to help ensure educational effectiveness in U.S. colleges and universities. That is to say, if a university or college wants to be accredited (or re-accredited), they must show the regional accreditation commission that faculty and administrators are engaging in, and supporting, outcomes assessment throughout the institution. Evidence of such activity demonstrates that the institution is attending to educational quality and effectiveness. For faculty at the program or departmental level, this requirement (depending on the accreditation commission) typically involves development of an assessment plan, a mission statement, student learning outcomes statements, curricular maps, rubrics, implementation of student learning data-collection methods, and reporting of assessment processes and uses to the institution. All regional accreditation commissions have incorporated assessment requirements into their accreditation review procedures over the last 15 years or so, and higher education in general has been experiencing the impact of those mandates ever since.

The requirement to undertake student learning outcomes assessment represents an important change in higher education accreditation processes and a new reality for college/university foreign language educators. This state of affairs, I would argue, is a traditional research concern in language education studies. Applied linguistics and language education research has historically taken an interest in the ‘impact’ of accountability regimes on U.S. language education. Most of this type of research, however, has been conducted at the K-12 level. For example, researchers have examined the impacts of No Child Left Behind legislation on English language learners in U.S. public schools, given the special provisions the law stipulates for such students. Researchers have also looked at the impacts of NCLB on foreign language

education in U.S. public schools, generally finding that the diversion of resources to math and language arts education enervates foreign language education in various ways (e.g., Rhodes & Pufhal, 2010; Rosenbusch & Jensen, 2005; von Zastrow & Janc, 2004). I likewise regard mandated outcomes assessment as a NCLB-type phenomenon in that it is an accountability system with intended national-level impacts, driven by political discourses concerned with educational quality, designed to reform educational delivery, and, thereby, educational effectiveness in U.S. higher education.

Given the pervasiveness of outcomes assessment, I noticed there had been comparatively little investigation into its influence on FL programs in higher education. The desired impact of assessment—from the accreditation commission point of view—is increased program-level inquiry into educational processes and use of assessment toward educational reform. In my view, it was unclear if this is happening in post-secondary FL education. Prior to my study, outcomes assessment in college FL programs had been addressed in a limited way in the FL research literature, mainly via commentary from a few foreign language assessment experts (Liskin-Gasparro, 1995; Norris, 2006, 2009) and experienced assessment practitioners (Byrnes, 2008), as well as in a few case studies describing experiences undertaking assessment in university FL programs (e.g., Liskin-Gasparro, 1995; Ramsay, 2009; Walther, 2009). The general view emerging from these nascent discourses was that the national outcomes assessment mandate—originating in accreditation standards and promulgated by colleges and universities—is not an especially useful form of programmatic inquiry and reform. This largely accurate observation notwithstanding, to my mind there had been no empirically-based, theoretically-driven understanding of how, whether, and why mandated outcomes assessment facilitates or hinders improvement of FL education in U.S. colleges and universities. I thus thought we needed to

know more about the role of accreditation-mandated assessment in post-secondary FL educational delivery.

***PLJ- What was the specific focus of your research?***

I decided to undertake an exploratory investigation of the general usefulness of assessment activity in FL higher education BA major programs, and, importantly, whether institutional assessment requirements, in particular (these for the most part the result of accreditation mandates), make those assessment activities more or less useful for FL educators.

Note that a key notion of interest was concrete assessment *use*—I wanted to investigate the ‘impact’ of mandated assessment in terms of how it seemed to contribute to real actions and decision-making addressing educational processes. The ultimate outcome of accreditation-mandated, assessment-based accountability systems is precisely this objective: concrete use of assessment findings (i.e., the information gleaned about student learning) toward program reform and improvement. Again, the FL research literature (and assessment research elsewhere) suggests that mandated assessment is not effective in accomplishing this aim. Rather, assessment requirements seem more successful at getting certain types of assessment infrastructure in place (e.g., a program mission/vision statement, student learning outcomes statements, curriculum maps, etc.), but less effective in “closing the assessment loop”—that is, having educators make decisions and take reformative actions on the basis of consistent, thoughtful consideration of student learning evidence. I was particularly interested, then, in knowing if the implementation of particular accreditation-driven assessment requirements were helping FL educators realize this specific goal. My overarching research question (RQ) captured this general interest: How does accreditation-mandated assessment contribute to assessment use in tertiary foreign language programs (RQ1)?

Further, to get a sense of how (and whether) assessment mandates engender useful assessment, I structured the study borrowing methods from program evaluation and organizational learning research. Overall, I wanted to know if there was a relationship between the types of assessment-related structures that an institution asks FL educators to put in place and the degree to which educators are using assessment. I thus wanted to collect information on (a) how much assessment *capacity* existed in FL programs—that is, the extant infrastructures, practices, personnel characteristics, and institutional context, factors, etc., that enable programs to do useful assessment; and (b), how much assessment *use* was happening in FL programs—or, the various actions, decisions, instances of learning, awareness-raising, etc. that result from engaging in assessment.

The reason I wanted to know about assessment capacity and use is because of a hypothesized relationship between the two: more assessment capacity is proposed to result in more assessment use (an idea I borrowed from program evaluation and organizational learning research). That is to say, if a program has a robust and strong ability to do assessment—for example, faculty collaborate and communicate productively together on assessment work, they receive ample assessment support and incentives, leadership is committed to assessment, and so on—there is a greater likelihood that assessment efforts will be productively useful and used in desired ways.

Moreover, and importantly, the aspects of FL program assessment capacity I particularly wanted to know about were the specific assessment requirements and recommendations issuing from institutional (i.e., accreditation) assessment mandates. During accreditation review, accreditation commissions will look for evidence that institutions are managing and supporting assessment in particular ways; for example, whether they are providing assessment expertise and

funding, or whether they are requiring programs to report on assessment activity (among other things). By exacting these requirements, accreditation commissions instantiate a type of assessment capacity in FL programs, and I was interested if, by meeting these requirements and developing capacity in this way, faculty were better able to undertake and use assessment toward improved processes.

Of course, the ability to undertake assessment/assessment capacity is comprised of a complex set of local conditions, some of which will be the result of responding to institutional, accreditation-related requirements and some the result of other forces (e.g., motivated faculty proactively building assessment capacity themselves). My aim, then, was to perform a kind of assessment ‘audit’ of as many college-level FL programs as I could, capturing each program’s overall level of assessment capacity and assessment use, and then look for any general relationships between assessment capacity and assessment use, but also the specific relationships between assessment use and the institutionally driven aspects of a program’s assessment capacity. Structuring the study in this way I thought would enable me to see what aspects of assessment capacity in a FL program would be most related to assessment use, and, especially, whether the institutionally mandated (and accreditation-driven) dimension of a program’s assessment capacity was more or less related to FL educators using assessment toward improvement-oriented ends. I posed four sub-research questions to capture these research concerns: What is the nature and extent of assessment capacity in college FL programs (RQ1.1)? Which assessment uses are occurring in college FL programs (RQ1.2)? What is the relationship between assessment capacity and assessment use in college FL programs (RQ1.3)? What is the relationship between accreditation-mandated assessment capacity and assessment use in college FL programs (RQ1.4)? As noted above, parsing the study into these four sub-questions was

meant to shed light on my overarching research question: How does accreditation-mandated assessment contribute to assessment use in tertiary foreign language programs (RQ1)?

***PLJ- How did you organize your survey tool and to whom did you send it?***

Again, I needed to know about two categories of assessment activity happening in foreign language programs: the amount of assessment capacity and the amount of assessment use. I designed an online questionnaire that would gather this type of information. The first part of the survey asked about assessment capacity. The respondent was asked to indicate the degree to which certain capacity elements existed in their program in nine categories: (a) institutional assessment support, (b) institutional assessment policies and governance patterns, (c) assessment infrastructures, (d) programmatic assessment support and resources, (e) assessment-related leadership, (f) assessment-conducive program ethos, (g) assessment collaboration, (h) assessment communication, and (i) high-quality assessment activities and conditions (a catalogue of “best-practices” demonstrated by highly competent, experienced programs with a history of successfully productive assessment). Each category had a number of questions (anywhere from 5 to 15) asking the respondent to rate the degree to which a particular aspect of assessment capacity existed in their program: for example, the extent to which “...people in the program are willing to collaborate with colleagues on outcomes assessment/program development work?” (1 = not at all, 2, 3, 4 = very much/a lot). Note that categories *a*, *b*, and *c* (institutional support, institutional policies and governance patterns, and assessment infrastructures) were comprised of items defined in terms of the particular requirements and recommendations that accreditation commissions stipulate, and thereby the sorts of practices and structures that institutions require faculty implement in their programs. Taken together, the capacity sub-categories were thought to comprise a relatively comprehensive assessment capacity framework, which, if present in a

college FL program (across all the categories), would strongly dispose educators to use assessment in a variety of meaningful ways.

The second part of the questionnaire asked about how assessment was being used in the program: for example, “To what extent has anyone in your department/program used outcomes assessment findings (i.e., information about student performance) to develop/modify instruction/courses?” (1 = not at all, 2, 3, 4 = very much/a lot). The assessment capacity and use questionnaire items, so designed (along with a number of background information items), enabled me to collect information on the degree of assessment capacity and assessment use in FL programs at U.S. universities.

The questionnaire was created online (using *SurveyMonkey*) and sent out to university and college FL educators via email. I targeted FL faculty who would be knowledgeable about assessment happening in their department or program. Typically these people were chairs, language section/division heads, or assessment specialists, though in many cases the most knowledgeable people were lower-division language coordinators or like-individuals.

Before sending out the questionnaire, I searched FL program websites and compiled a list of all university FL major programs in the U.S. (to the extent possible), as well as a list of contact names and email addresses for an assessment informant at each program. I contacted approximately 1513 FL major programs via email with a web link to the questionnaire. In the end, I collected data from 100 FL major programs.

Respondents from the 100 programs provided information about the degree of assessment capacity and use associated with their major program(s). Once the data were in, I wanted to find out which of the nine capacity categories in FL programs were most strongly related to assessment use (using inferential statistics), and I was particularly interested in the strength of

relationship between the more ‘institutionally-oriented’ capacity categories and use. Again, I wanted to know whether more or less assessment capacity would correlate with more or less assessment use, and particularly if the existence and amount of institutionally-driven assessment capacity would enable a program to use assessment in constructive ways.

***PLJ- Generally speaking, what were your findings?***

At this point I should point out that the number of programs represented in the study ( $n = 100$ ; 6.6% of the population) does not generalize well to all FL programs at U.S. universities, which was one of the goals of the study. Yet, there was wide variability in the results, which I take to be representative of the *types* of assessment capacity/use trends in the U.S. higher education FL program population rather than the precise proportions in which these trends occur.

The overall conclusions to be drawn from my research results are that (a) programs differ greatly in amounts and types of assessment capacity and assessment use; (b) key types of capacity are more strongly predictive of assessment use than others; and (c) accreditation-related, institutionally-driven types of assessment capacity do not have a strong relationship with assessment use.

Elaborating on these findings further (by research question), the existence of assessment capacity (RQ1.1) and assessment use (RQ1.2) in FL programs indicates some interesting trends. Regarding the extent of assessment capacity (RQ1.1), results suggest such capacity is highly variable with some programs reporting especially high amounts of capacity in all the nine-capacity areas, and others reporting relatively little. On balance, however, most programs have nascent (or entirely absent) assessment capacity in a majority of the capacity areas, *except* in the institutionally-driven categories for which a greater majority of programs reported fully developed or developing capacity levels.

Indications of the extent of assessment use in FL programs (RQ1.2) are similar. Some programs are frequently using assessment in a variety of ways, though the greater majority reported little assessment use. Moreover, a problematic finding for accreditation commissions, arguably, is that the most frequently noted type of assessment use is to meet institutional and accreditation requirements (e.g., reporting on assessment to the administration).

Given these levels of assessment activity, a number of inferential statistical procedures were conducted to find relationships between the various capacity categories and assessment use (addressing RQ1.3 and RQ1.4). Results addressing overall relationships between capacity and use (RQ1.3) suggest that the capacity types most related to, or predictive of, assessment use are associated with certain personnel-related factors within programs, particularly assessment-conducive program ethos (i.e., attitudes/thinking), effective assessment collaboration, effective assessment communication, and high-quality assessment activities/conditions. A second group of capacity categories—which include program assessment leadership, program assessment support, and assessment infrastructures—have weaker and moderate relationships with assessment use.

Moreover, the types of capacity least (and most weakly) related to assessment use are, crucially, the institutionally-driven capacity varieties, institutional support and institutional policies/governance patterns in particular. I interpreted these results as evidence addressing RQ1.4: that is, there appears to be a weak relationship between the accreditation-driven, institutionally mandated assessment capacity and assessment use in post-secondary FL programs. Further evidence to support this claim was found by dividing programs into high-use and low-use groups and analyzing levels of capacity by group. Programs that demonstrate little assessment use ( $n = 66$ ) tended to indicate the *most* capacity in the institutionally-driven assessment capacity categories: institutional support, institutional policies/governance patterns,

and assessment infrastructures. By contrast, programs that demonstrate high amounts of assessment use ( $n = 34$ ) indicated the *least* amount of capacity in the institutionally-driven assessment capacity areas and greater amounts of capacity in the more ‘program-internal’ capacity areas, particularly assessment collaboration, assessment communication, and high-quality assessment activities/conditions.

The overarching conclusion I drew from these findings is that likelihood of a group of FL educators using assessment productively is more related to factors *originating within* programs—the way people think and act, and the methodological quality of their assessment efforts—than assessment mandates, requirements, and support coming from an institution. A somewhat grim implication of these findings for accreditors and university administrations, I argue, is that the current assessment-based accountability system—spearheaded by accreditation commissions and operationalized via institutional accreditation review—does not seem to be a sufficient condition for useful assessment to occur in FL programs. Certainly, institutional requirements and assessment support will help assessment efforts along and to some degree create circumstances that make assessment use possible. And it can be fairly said that without accreditation requirements spurring assessment on, very little assessment would be happening in post-secondary FL programs at all. Never-the-less, my results indicate that in FL programs where assessment capacity is strongly defined by a compliance with institutional requirements, with little evidence of additional program-internal types of assessment capacity (conducive leadership, assessment ethos, collaboration, communication, etc.), little meaningful assessment use is likely to ensue.

***PLJ- Where and when will the complete results be available?***

A full discussion of the results and findings will appear in a volume on assessment in FL education to be published by the National Foreign Language Resource Center and edited by John Norris, Yukiko Watanabe, and myself. It will be available sometime in 2013.

***PLJ- Please talk about the usefulness of evaluation/assessment to FL education.***

Despite generally negative feelings toward required assessment throughout higher education, I would argue that college language educators have much to gain from implementing assessment (or evaluation<sup>2</sup>) in their programs. One of the unsurprising findings from my study is that assessment in college FL education continues to be a largely bureaucratic exercise in satisfying administrative requirements. Of course, assessment is envisioned to be much more. Assessment is meant to be, rather, a process of thoughtful, systematic inquiry into the delivery of education in college programs (Norris, 2006; Byrnes, Maxim, & Norris, 2010). Indeed, there is clear evidence that assessment can function in just this way and that language educators use assessment as a tool to understand and reform the language learning and teaching happening within their programs. To find out more about such examples, I encourage a visit to the FLPEP website<sup>3</sup> and perusal of the various publications and presentation listed there to see how FL educators have been using assessment and evaluation to deepen their understanding of what they do as language teaching professionals (I particularly recommend Norris, Davis, Sinicrope, & Watanabe, 2009; Byrnes, Maxim, & Norris, 2010; see also the forthcoming NFLRC volume, noted above, for additional examples).

---

<sup>2</sup> My response specifically addresses the benefits of *assessment* in higher foreign language education (rather than evaluation) as per the focus of my dissertation research.

<sup>3</sup> <http://nflrc.hawaii.edu/evaluation/index.htm>

Persuasive arguments have been made elsewhere about the usefulness of assessment in FL programs (see in particular Byrnes, 2008, and Norris, 2006). To make the case here, I put to readers a preeminent example of assessment's potential: the Georgetown University German Department (GUGD), an exemplary case of FL educators using assessment to bring about an extensive curricular transformation of a German major program (described in detail in Byrnes, Maxim, & Norris, 2010). Assessment (and evaluation) was used at GUGD to address a known issue in university language education programs: creating a coherent major curriculum articulating beginning first-year instruction through to advanced-ability levels. In attempt to bridge the divide between advanced, literary types of language performances, and beginning, communication-oriented language learning, GUGD faculty and staff devised a specific approach to language pedagogy (focusing specifically on writing abilities) that used conceptions of language learning grounded in systemic functional linguistics, these codified in specific articulations of German writing literacy, and realized pedagogically via a curriculum organized around sequenced, genre-defined writing tasks.

Assessment was fundamental to this curricular innovation. Early on in program development efforts, assessment was understood as part of a symbiotic relationship with curriculum and instruction, each meant to inform and reform the other. It will be important to point out that the role of assessment in this transformation—crucially—was not as a post hoc, auxiliary add-on arising out of institutional requirements. Rather, assessment was seen as an integral and necessary part of the innovation process. Fundamentally, assessment provided the window through which personnel could “see” their curricular innovations in action, and as such, information about student learning became a bellwether of curricular functioning. Over time and after many assessment-driven reform iterations, feedback from student writing performances

(and other sources) was continually fed back into curricular modifications, the result being a carefully crafted, integrated major curriculum moving in a coherent trajectory from beginning to advanced levels.

For language educators delivering instruction in less commonly taught languages, there is much to recommend for assessment (and evaluation) on the basis of the GUGD example. A remark I heard during an assessment workshop captures well assessment's potential: "knowledge is power." Ultimately, I think, assessment in the GUGD instance empowered people to take control of the educational processes happening in their programs and refashion them in ways reflecting the educational vision and values of educators in that particular context. To further describe some of the GUGD's assessment achievements, and the achievements of others elsewhere, I will end with a shortlist of key transformative possibilities (elaborating on a list from Norris & Liskin-Gasparro, 2009, as well as from Byrnes, Maxim, & Norris, 2010) that can occur when people energetically engage in a process of educational inquiry using assessment and evaluation methods.

By engaging in assessment, FL educators are able to...

- specify valued expectations for FL development and achievement, and provide clear learning and instructional targets for teachers and students;
- concretely evidence students' actual knowledge, abilities, and dispositions to various audiences;
- enable systematic reform of curriculum and instruction on the basis of evidence;
- align instruction with curriculum
- initiate fruitful, illuminating dialogues within and across FL programs and constituencies

- create productive, collaborative communities (potentially across traditional disciplinary divides) focusing collectively on student learning;
- raise awareness about noteworthy aspects of a FL program (e.g., learning targets, student achievement) among students, faculty, and the institution.
- engender programmatic thinking (e.g., in terms of how course content/instruction articulates with proximate courses, as well the overall curricular structure);
- streamline uses of assessment tools, revealing inefficiencies and redundancies;
- cogently support requests for resources
- demonstrate the value of FL studies to students, the institution, and the public (at a time when language education is increasingly asked to justify itself)
- increase program ownership;
- beget cycles of iterative improvement;
- expand understandings of language learning and pedagogy, furthering scholarship and professional development;
- improve student learning in targeted, systematic, locally-valued ways

***PLJ- What would you suggest to faculty as a first step to begin using assessment/evaluation in their Portuguese programs?***

It will be important to first point out that advice for starting assessment or evaluation efforts will vary depending on the nature of the program (e.g., the size of the program), the program's purposes for undertaking assessment, and, more generally, the social and cultural context within which the program functions. Further, much could be said about starting assessment/evaluation (e.g., gauging evaluation/assessment 'readiness') since gearing up and

planning to do assessment is itself a crucial first stage in laying the groundwork for useful, meaningful assessment/evaluation to come. Never-the-less, a few recommendations can be made to help initial efforts off to a good start, which would seem to apply to most college FL programs (in my opinion).

*Be inclusive in initial efforts.* The usefulness and productiveness of assessment/evaluation is increased when it proceeds democratically, with everyone ‘on board’, and reflecting the desires and goals of a program majority (ideally). During initial discussions, then, in which assessment is being planned or pitched, try to be as inclusive as possible of all the relevant stakeholders or stakeholder groups in your program (if practical do so). A few key groups to keep in mind: people in leadership roles since these individuals will be instrumental in supporting assessment/evaluation efforts and using assessment results in decision-making. Also, try to include students. Students bring a useful perspective to assessment efforts; they can speak to assessment needs and respond well to involvement in educational decision-making that ultimately impacts them most<sup>4</sup>. Note, however, that the above advice depends on people being willing and generally open to (or at least not overtly obstructive of) the idea of assessment/evaluation. Contextual factors may contradict the above; it can be better to move forward with a minority of willing participants in the interest of getting assessment/evaluation done, rather than persist with a reluctant, larger group.

*Get technical assistance.* The entire assessment/evaluation enterprise is helped along considerably with some expert advice. To be sure, the bulk of assessment/evaluation planning, methods development, data collection, interpretation, etc. will need to be done by non-experts/local people, but it can be productive to get some expert assistance, especially at beginning stages. Such expertise can be sought from a number of sources. Obviously, anyone

---

<sup>4</sup> Assessment is also good professional training for graduate students.

within the program with assessment, evaluation, or even language testing expertise should be involved in development projects. Often, there will be some kind of assessment support unit at your institution that exists to help faculty and staff with their assessment/evaluation work. Look into what this entity offers. Most provide workshops and training on assessment methods, or consulting services, even technology support. Assistance may also be found in other departments and units at your institution. Are any other language departments or units doing assessment? Is there assessment/evaluation research expertise in a local education department or departments/programs elsewhere? Search the institution thoroughly for potential support since there may be conveniently local expertise able to lend a hand.

*Self-educate.* Ultimately, those undertaking assessment/evaluation will take on much of the burden of assessment/evaluation work themselves, and to do so, some assessment/evaluation skills and knowledge will be required. To the extent possible, it will be worth having participants read up on some evaluation/assessment literature, as well as participate in any workshops and hands-on training opportunities (see above). As a first step toward doing so, I particularly recommend looking into the various resources offered by FLPEP at the project website (see above). These resources are specifically designed for foreign language assessment/evaluation and foreign language educators. As a start, and to get acquainted with some key ideas, I suggest looking at a document entitled *Roles and responsibilities for evaluation in foreign language programs* (Norris & Watanabe, 2007<sup>5</sup>). This is a short, step-by-step guide to the particular evaluation approach espoused by the FLPEP group (one informed by utilization-focused ideas developed by Michael Quinn Patton) and provides some of the basic ideas and overall FLPEP framework for doing and using evaluation and assessment in university/college foreign language programs. The other publication I would highly recommend is *Toward useful program evaluation*

---

<sup>5</sup> <http://nflrc.hawaii.edu/evaluation/files/Roles%20and%20Responsibilities%20booklet.pdf>.

*in college foreign language education* (Norris, Davis, Sinicrope, & Watanabe, 2009). This volume presents a number of case studies depicting university FL educators' experiences using evaluation/assessment toward program development. It provides useful examples of how evaluation can be done in FL programs and vividly demonstrates the possibilities for FL program understanding and reform using assessment/evaluation methods. In addition to these two publications, there are numerous like resources at the site that will help FL educators build practical assessment/evaluation skills and knowledge, as well as show how assessment and evaluation is being used to good effect in college FL programs.

*Seek out funding.* Following on from points above about resources, securing extra funding for assessment will make some of the above recommendations possible. For example, funding can be used to bring in assessment experts/consultants, for assessment-related conference travel, for professional memberships, for faculty course buy-outs, to pay graduate assistants for assessment work, to purchase equipment (e.g., recording devices, etc.), software, reference materials, and so on. Despite the fact that all of higher education is somewhat cash-strapped currently, you may find that assessment funding is still available. Entreaties should be made locally, to whoever holds the purse strings in the department/program, to set aside funds for assessment/evaluation work. Further, your institution may provide funding for developing assessment capacity. Some institutions provide assessment grants and other forms of financial support, and I have seen people unaware such resources exist. Moreover, bear in mind that competently prosecuted, productive assessment/evaluation helps with subsequent requests for funding, and it may be possible to parlay successful assessment efforts in additional money some time down the road.

*Start small.* Finally, I would tentatively recommend a small, circumscribed assessment/evaluation project to get things started (if possible). I might be hard pressed to define what a “small” assessment project is, though something like a preliminary investigation into student perceptions (e.g., of a course, a sequence of courses, or their own learning progress), I imagine, would be of the scope I have in mind. It is important that the first assessment/evaluation effort be a successful one—to that end, a modest project has a number of attractive aspects. First, a useful assessment project is necessarily a finished one, and a small, feasible project may be easier to get accomplished in a timely way. Further, a smaller project might require less in the way of resources, which will be helpful if such resources are unavailable or scarce. Also, a modest, initial assessment success over the short term will demonstrate the usefulness and efficacy of assessment/evaluation, potentially generate support for future projects, and possibly win over those who lacked enthusiasm the first time round. Ultimately, however, the size and scope of an assessment or evaluation project will be determined by the particular needs and purposes of local stakeholders. The above advice, then, is simply to point out that feasibility and the likelihood of really getting a project done should be kept in mind when starting out with your program’s assessment/evaluation endeavors.

## References

- Byrnes, H. (2008). Owning up to ownership of foreign language program outcomes assessment. *ADFL Bulletin*, 39(2-3), 28-30.
- Byrnes, H., Maxim, H., & Norris, J. M. (2010). Realizing advanced FL writing development in collegiate education: Curricular design, pedagogy, assessment. *Modern Language Journal*, Monograph. Cambridge: Blackwell
- Liskin-Gasparro, J. (1995). Practical approaches to outcomes assessment: The undergraduate major in foreign languages and literatures. *ADFL Bulletin*, 26(2), 21-27.
- Norris, J. M. (2006). The why (and how) of assessing student learning outcomes in college foreign language programs. *Modern Language Journal*, 90(4), 590-597.
- Norris, J. M. (2009). Introduction to the volume. In J. M. Norris, J. McE. Davis, C. Sinicrope, & Y. Watanabe (Eds.), *Toward useful program evaluation in college foreign language education* (pp. 1–4). Honolulu, HI: University of Hawai‘i, National Foreign Language Resource Center.
- Norris, J. M., Davis, J. McE., Sinicrope, C., & Watanabe Y. (Eds.) (2009). *Toward useful program evaluation in college foreign language education*. Honolulu, HI: University of Hawai‘i, National Foreign Language Resource Center.
- Norris, J. M., & Liskin-Gasparro, J. (2009). The consequences of accreditation and outcomes assessment for college foreign language programs. Paper presented at the invited colloquium “Critical issues at the interface between assessment and U.S. language education policy”, at the annual conference of the American Association of Applied Linguistics, Denver, CO.

- Norris, J. M., & Watanabe, Y. W. (2007). Roles and responsibilities for evaluation in foreign language programs. Retrieved from <http://www.nflrc.hawaii.edu/evaluation/files/Roles%20and%20Responsibilities%20booklet.pdf>
- Ramsay, V. (2009). Study abroad and evaluation: Critical changes to enhance linguistic and cultural growth. In J. M. Norris, J. McE. Davis, C. Sinicrope, & Y. Watanabe (Eds.), *Toward useful program evaluation in college foreign language education* (pp. 163–182). Honolulu, HI: University of Hawai‘i, National Foreign Language Resource Center.
- Rhodes, N. C., & Pufahl, I. (2010). *Foreign language teaching in U.S. schools: Results of a national survey*. Washington, DC: Center for Applied Linguistics
- Rosenbusch, M. H., & Jensen, J. (2005). Status of foreign language programs in the NECTFL states. *NECTFL Review*, 56, 26-37.
- von Zastrow, C., & Janc, H. (2004) *Academic atrophy: The condition of the liberal arts in America’s public schools*. Washington, DC: Council for Basic Education.
- Walther, I. C. (2009). Developing and implementing an evaluation of the foreign language requirement at Duke University. In J. M. Norris, J. McE. Davis, C. Sinicrope, & Y. Watanabe (Eds.), *Toward useful program evaluation in college foreign language education* (pp. 117–138). Honolulu, HI: University of Hawai‘i, National Foreign Language Resource Center.