CrossMark

Lang. Teach. (2016), **49.4**, 592–595 © Cambridge University Press 2016
doi:10.1017/S0261444816000215

Research in the News

The Language Flagship Proficiency Initiative

The Center for Language Teaching Advancement (CeLTA) at Michigan State University (MSU) was awarded a U.S. federally funded two-year grant in 2014 to investigate the status of proficiency, and the promotion of proficiency, in foreign language programs. The grant was awarded to Paula Winke and Susan Gass, as principal investigators. The purpose of this initiative was to assess language proficiency development and to show the impact that introducing assessment practices into established language programs has on pedagogical practices and, ultimately, on proficiency outcomes. In total, three grants were awarded, one to MSU, one to the University of Utah, and one to the University of Minnesota. The overall goals were the same, although each institution approached the task in slightly different ways and with slightly different emphases.

At MSU, we focused our attention on Chinese, French, Russian, and Spanish. We formulated four main goals for these foreign language programs: (1) monitor students' proficiency gains, (2) identify internal and external variables that contribute to those gains, (3) work with faculty and instructors to provide them with the training necessary to promote clear, articulated proficiency goals and student proficiency gains, and (4) create an environment where proficiency assessment is the norm.

We created a research team that has involved faculty members (Charlene Polio, Dan Reed, Koen Van Gorp), students (Jessica Fox, Shinhye Lee, Suzie Johnston) and various others who occupy leadership positions in those foreign languages that we were directly investigating. With the assistance of our graduate students, our Project Manager (Elizabeth Webster), and the Executive Associate Director of CeLTA (Angelika Kraemer), we began testing students in the Fall of 2014 with the initial goal of understanding the proficiency levels of our students based on ACTFL proficiency assessments[1] for reading, listening, and speaking (respectively the Reading Proficiency Test (RPT), the Listening Proficiency test (LPT) and the Oral Proficiency Interview – Computer (OPIc), as offered by Language Testing International). These computer-based proficiency tests are standardized assessments evaluating the Novice to Distinguished proficiency levels following the ACTFL scale, although our assessments only went to Superior.

We administered more than 5,500 tests during the first academic year (Fall 2014 and Spring 2015) and more than 5,100 tests during the second academic year (Fall 2015 and Spring 2016) across four levels of study (See Table 1). We did not test first semester first year students (for example, Chinese 101) given our belief that there was little 'proficiency' at that

[1] See: www.languagetesting.com/actfl-proficiency-scale

https://doi.org/10.1017/S0261444816000215 Published online by Cambridge University Press

Table 1 Number of test administrations across four semesters

Proficiency test	Fall 2014	Spring 2015	Fall 2015	Spring 2016
OPIc	1,028	1,010	671	1,179
RPT	933	825	601	1,054
LPT	927	812	588	1,017

OPIc = Oral Proficiency Interview – computer; RPT = Reading Proficiency Test; LPT = Listening Proficiency Test

Table 2 Distribution of number of test administrations per languages across semesters

Proficiency test	Fall 2014	Spring 2015	Fall 2015	Spring 2016
Spanish	1,805	1,357	917	1,734
French	605	811	486	961
Chinese	340	354	342	343
Russian	138	125	115	212
Total	**2,888**	**2,647**	**1,860**	**3,250**

level of coursework. The highest number of tests was administered in the Spanish language program, the lowest number in the Russian language program (See Table 2).

The results of this large-scale testing program have led to beneficial in-house discussions with program directors and instructors of the different languages involved. Meeting at least once a semester, we discussed the impact of proficiency testing on the (innovation of the) language curriculum and the setting of proficiency goals for the different course levels. For example, we found that for a particular language, most students in the 2nd year were at the Novice High – Intermediate Low proficiency level for speaking, listening and reading, at the Intermediate Mid – High level in the 3rd year, and at the Intermediate Mid – High level in the 4th year. For some languages we found that the proficiency level of reading or listening might be constantly higher or lower than the speaking level. These results caused program directors and instructors to think about the desirability of these score assignments and how they relate to the content of the language curriculum.

In addition to the large number of test results, we administered an in-depth background questionnaire that focused on study abroad, the use of language outside of the classroom, technology use, goals for language learning, and so on. These background variables will allow us to break down the overall findings for specific types of learners (e.g., study abroad or not) and look at what factors affect what language skills the most, not only cross-sectionally, but also longitudinally.

Another dataset came from focus groups, conducted by Meg Malone from the Center for Applied Linguistics. We conducted eight focus group interviews with students (n = 75), three of which were focus groups with study abroad students (n = 12), and six focus groups with instructors (n = 56). These qualitative data have been coded primarily by our graduate

https://doi.org/10.1017/S0261444816000215 Published online by Cambridge University Press

students (Fox and Lee) using NVivo. All the data in NVivo were tagged with demographic information: student, instructor, language, gender, age, and level of instruction. We followed an open-coding methodology (Dörnyei 2007) to identify the major themes that emerged in the qualitative data and we allowed the data to shape emergent themes. Three initial nodes (goals, assessments, pedagogy) emerged and we further subcategorized them as follows: Goals (program, personal), Assessments (type of assessments in class, standardized (ACTFL) proficiency tests, placement tests), Pedagogy (level, study abroad). This dataset allows us to investigate how proficiency goals, development, and teaching are connected for instructors and students, and how perceptions and reported actions for both groups relate to one another. One of the interesting findings from the first focus groups is how students' understanding of language proficiency is intertwined with the curriculum (i.e., grammar or literature), its perceived goals (e.g., reading and writing in the upper levels) and limitations (e.g., becoming fluent is difficult in a classroom setting), and their personal goals (e.g., to be able to converse with native speakers). Moreover, students view proficiency especially as speaking proficiency and have a tendency to separate it from (their idea of) fluency: 'I think PROFICIENCY is more like EDUCATIONAL kind of and FLUENCY is like you're BORN with this. I'm not sure.' (French student). Most instructors on the other hand have a clear understanding of proficiency ('can do' versus 'know how to use') and proficiency development (e.g., the ACTFL proficiency levels and scales). Most instructors adhere to a communicative philosophy (e.g., get students to 'communicate coherently enough that they can be understood by native speakers'), but feel often constrained by the 'old' or 'only available' textbooks and materials in the programs: 'I'm trying to put something together, instructional material that can teach the students.' (Chinese instructor). As such, the ongoing analyses will provide us with an in-depth view on foreign language proficiency teaching and learning at MSU, a university which is typical of many in the United States.

This view on teaching and learning will be refined as we incorporate data from classroom observations that were supervised by Charlene Polio. The main focus of these observations was to collect data concerning the extent to which proficiency is the focus of the courses at the different levels. A total of twelve classes taught by six teachers were observed: Three instructors, each teaching two Chinese, French, and Spanish 100 level classes respectively and three instructors, each teaching two Chinese, French, and Russian 300/400 level classes. Each classroom observation was documented by two observers, independently of one another, using a structured observation scheme detailing content or focus of the activity, participant organization, student language modality or skill focus, materials and visuals used, and amount and functionality of English used in the classroom. The data of these classroom observations will be linked to both teacher and student reports of classroom activity in the focus groups and the results on the different proficiency tests. This will allow us to triangulate the data and gain insight into the amount and impact of proficiency-based teaching in language courses at MSU.

We are in the process of further analyzing our data with some of our graduate students and researchers external to MSU. We are investigating (a) the impact of study abroad on language acquisition and how programs can best reintegrate study abroad students into the language program, (b) self-assessment and how it is related to measures of proficiency, (c), the impact of technology use on language acquisition, (d) the attainment of proficiency across non-majors, minors, and majors, and (e) structural equation models of proficiency (which

https://doi.org/10.1017/S0261444816000215 Published online by Cambridge University Press

may show us which independent and individual-difference variables account for differences in proficiency attainment).

We anticipate that numerous projects will continue to branch off of the original data. As an example, we illustrate this through a new project that the two Principal Investigators (Winke and Gass) are undertaking with Bill VanPatten and Koen Van Gorp. From our testing project, we asked ourselves how students perceive the assessments (some were tested multiple times over the two-year period) and whether the students really performed at their best. Could incentivization make them perform better (that is, care more and thus work harder during testing)? We questioned the motivation that students have for achieving any particular level of ability. In many university settings, the only requirement for language students is seat time (i.e., X number of credit hours completed). So we asked ourselves: What would happen if we motivated students by giving them an incentive? What if we demanded that meeting the pre-established proficiency goals on the proficiency tests would contribute to their final grade, for example for 5 or 10% of their final grade? Would this work as incentivization and result in more students reaching the program goals?

We are currently looking into the effects of different forms of incentivization with the students in second year Spanish courses at MSU (n = 217) and with a smaller group of Chinese language learners (n = 97). If the incentivized groups outperform comparison groups (who have no incentivization), then we may need to suggest that programs should only enact proficiency-based exit-test if there are requirements tied to them that articulate the programs' stated goals (see Chalhoub-Deville 1997). We would further suggest that exit-test requirements be put in place, but such tests will have to be integrated within the program and count as part of the students' program requirements.

References

Chalhoub-Deville, M. (1997). The Minnesota Articulation Project and its proficiency-based assessments. *Foreign Language Annals* 30, 492–502.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.

Susan Gass
gass@msu.edu

Paula Winke
winke@msu.edu

Koen Van Gorp
vangorpk@msu.edu

https://doi.org/10.1017/S0261444816000215 Published online by Cambridge University Press