AAUSC 2014 Volume – Issues in Language Program Direction

Innovation and Accountability in Language Program Evaluation

John Norris, Georgetown University
Nicole Mills, Harvard University
Editors

Stacey Katz Bourns, Harvard University
Series Editor

CENGAGE Learning®

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

CENGAGE Learning®

AAUSC 2014 Volume - Issues in Language Program Direction: Innovation and Accountability in Language Program Evaluation
John Norris, Nicole Mills, and Stacey Katz Bourns

Product Director: Beth Kramer

Product Assistant: Jacob Schott

Marketing Brand Manager: Christine Sosa

IP Analyst: Jessica Elias

IP Project Manager: Farah Fard

Manufacturing Planner: Betsy Donaghey

Art and Design Direction, Production Management, and Composition: Lumina Datamatics, Inc.

© 2016, Cengage Learning

WCN: 01-100-101

ALL RIGHTS RESERVED. No part of this work covered by the copyright herein may be reproduced, transmitted, stored, or used in any form or by any means graphic, electronic, or mechanical, including but not limited to photocopying, recording, scanning, digitizing, taping, web distribution, information networks, or information storage and retrieval systems, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the publisher.

For product information and technology assistance, contact us at **Cengage Learning Customer & Sales Support, 1-800-354-9706**

For permission to use material from this text or product, submit all requests online at **www.cengage.com/permissions**
Further permissions questions can be e-mailed to **permissionrequest@cengage.com**

Library of Congress Control Number: 2014948811

ISBN-13: 978-1-305-27509-6

Cengage Learning
20 Channel Center Street
Boston, MA
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at **international.cengage.com/region.**

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

For your course and learning solutions, visit **www.cengage.com.**

Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com.**

Instructors: Please visit **login.cengage.com** and log in to access instructor-specific resources.

Printed in the United States of America
Print Number: 01 Print Year: 2014

Chapter 4

Designing an Embedded Outcomes Assessment for Spanish Majors: Literary Interpretation and Analysis

Judith E. Liskin-Gasparro and Raychel Vasseur, University of Iowa

Introduction

The contributions to this volume highlight different aspects of and approaches to foreign language program evaluation. Assessing student learning across an entire program is central to understanding a program's strengths and weaknesses, and it is also a starting point for principled and strategic actions to pursue improvement. The assessment of student learning outcomes (SLOs) comprises two stages, each complex in its own right: (a) designing and conducting the assessment and (b) making use of the results of the assessment to improve the program. The case study reported in this chapter deals with the first stage: (a) designing and implementing an assessment of the knowledge and skills in literary interpretation and (b) analysis of graduating Spanish majors at the University of Iowa.

Following a university-wide mandate in 2006 that all academic departments at the University of Iowa develop SLOs for their undergraduate majors and conduct outcomes assessments "both to help maintain program excellence and to prioritize areas for program development" (University of Iowa Outcomes Assessment, n.d.), the Department of Spanish and Portuguese at the University of Iowa has been engaged in developing SLOs, designing assessments, and gathering information on student learning. The SLOs, which were approved by the departmental faculty in May 2007, cover a range of areas in the Spanish major: content knowledge, speaking skills, and intercultural competence. This chapter reports on the process of developing a rubric to assess the content knowledge of graduating Spanish majors as demonstrated in essays written in advanced upper-division courses in Spanish or Spanish American literature.

The SLO for content knowledge is as follows: "ability to analyze, synthesize, and effectively present written information and argumentation in Spanish" (Department of Spanish & Portuguese, 2007). This overarching statement applies to the areas in which students take courses in the major: literature, culture/civilization, linguistics, and film. The SLO for content knowledge in the area of literature continues as follows: "Students produce extended interpretive and analytical essays on Hispanic literary texts, using scholarly sources to support their arguments" (Department of Spanish & Portuguese, 2007).

This chapter is structured as a case study, in which the authors—a faculty member in the Department of Spanish and Portuguese at the University of Iowa and a doctoral student in Second Language Acquisition at the same university—highlight the challenges of conducting a utilization-focused assessment (Patton, 2008) and the process of engaging stakeholders (i.e., departmental faculty) in developing and implementing the assessment. The authors constituted the outcomes assessment team for this project. To report on this case study and describe the process of conducting a utilization-focused assessment, this chapter begins with background on the project and the SLOs at the University of Iowa. Following this background information, we discuss Patton's (2008) utilization-focused evaluation framework and then describe the development of the assessment rubric and report on its use in an operational assessment. In the discussion, we consider the impact of the assessment on the curriculum in the Spanish major and implications of this project for language program directors (LPDs). The second stage of the outcomes assessment, which lies beyond the scope of this chapter but is central to effective program evaluation, will be to use the results of the assessment to enhance student learning by designing major projects that correspond better to students' interests and goals, increasing students' collaboration in their learning by sharing with them the rubrics designed for this project, and incorporating the learning outcomes for literary interpretation and analysis throughout the undergraduate Spanish program. The departmental faculty should reconsider—and will most likely decide to revise—the SLO related to content knowledge in Hispanic literature, as well as the requirement of 15 pages of academic writing in the most advanced courses in literature that students take to fulfill their major requirements.

Project Background

In this section we provide the institutional and educational context for the current outcomes assessment initiative, with a brief overview of the Spanish major at the University of Iowa, the SLOs for graduating Spanish majors approved by departmental faculty in spring 2007, and a brief account of the assessment work related to the SLOs for speaking skills and development of intercultural competence.

The Spanish Major at the University of Iowa

As of July 2014, there are 262 declared Spanish majors at the University of Iowa. The majority of these students (62%) are double majors, combining Spanish with another discipline. Although we do not have data on majors' career goals, anecdotal accounts indicate that most of the double majors view Spanish as an auxiliary to their other major, with the goal of expanding their professional opportunities in the United States or abroad.

The curriculum of the major is quite flexible. It consists of 12 courses, or 36 semester hours, starting with courses at the fifth-semester (post–general education requirement for graduation) level. Of these 12 courses, students must

take at least one course in each of four core areas: peninsular Spanish literature, Spanish American literature, linguistics, and culture (either continent). The other eight courses may be additional courses in the four areas, as well as coursework in creative writing or in Spanish language skills, including language for special purposes.

The courses in the major are grouped into three difficulty levels. The entry-level (2000-level) bridge courses include offerings in language skills (writing and speaking), as well as introductory courses in culture and literature. Students may count no more than four 2000-level courses for the major; they typically take two or three before enrolling in courses at the mid level of difficulty.

The many mid-level (3000-level) courses cover the usual range of curricular offerings in literature (genres, periods, and regions), culture (history, regions, or major cities), and linguistics (overview of Hispanic linguistics, sound structure, sociolinguistics, applied linguistics, and acquisition). Courses in literature, culture, and linguistics that students take during study abroad are most often at this level.

The highest level (4000-level) courses, of which students must take at least three, include specialized offerings in the three major curricular areas. By faculty agreement, each of these courses has a common requirement of at least 15 pages of formal academic writing; in addition, only 4000-level courses taken on campus (i.e., not courses taken during study abroad) fulfill this requirement for the major. Because they are the most advanced courses in the major, students typically take them during their last two semesters before graduation; they are also the courses that served as the site for the development of the outcomes assessment protocol for literary interpretation and analysis, which is the subject of this chapter.

Student Learning Outcomes for Graduating Spanish Majors

As mentioned earlier, the faculty of the Department of Spanish and Portuguese approved SLOs for graduating Spanish majors in spring 2007 in three areas: speaking skills, intercultural competence, and content knowledge. The SLOs in their entirety are displayed in Appendix 1 and are summarized in Table 4-1.

Table 4-1 Summary of SLOs for Graduating Spanish Majors

Learning area	Summary of SLO
Language skills	• Speaking (goal: Advanced Low) • Academic writing (cf. assessed via *knowledge* SLO)
Cultural dispositions	• Knowledge of cultural practices and products • Growth in cross-cultural attitudes, perceptions, behaviors
Content knowledge: literature, culture, linguistics	• Ability to analyze, interpret, and synthesize texts • Ability to construct an argument

The departmental outcomes assessment committee has been working on the development and implementation of the various aspects of the project since the SLOs were approved. In the first phase of the project, the Spanish Speaking Test (Center for Applied Linguistics, 1995) was used for the assessment of speaking ability, and faculty members and graduate students were trained in scoring procedures (Liskin-Gasparro & Leonard, 2011). In the second phase, an open-ended survey instrument was developed to assess students' perceptions of their culture learning and their growth in intercultural competence, as acquired in their academic courses and/or outside-of-class experiences in study abroad and in the United States (e.g., jobs, friendships) (Liskin-Gasparro & Leonard, 2012). In the most recent phase of the project, described in this chapter, we report on the assessment of content knowledge that Spanish majors acquire as the result of their undergraduate coursework in Hispanic literatures. We continue this case study report with the questions that we used to guide this project.

Guiding Questions

At this stage in the outcomes assessment project for the Department of Spanish and Portuguese, the following questions guided the development of the rubric to assess Spanish majors' skills in literary analysis and interpretation, as well as our analysis of and reflection on the development process.

1. Framed by a utilization-focused approach (Patton, 2008), what are the steps in developing a rubric to assess the ability of graduating Spanish majors at the University of Iowa to analyze and interpret literary texts?
2. What are the specific features of literary interpretation and analysis expected of a graduating Spanish major at the University of Iowa?
3. What are the challenges of using a utilization-focused approach for the assessment of learning outcomes in the Department of Spanish and Portuguese at the University of Iowa?

To summarize, the task we set for ourselves was the design and implementation of an embedded assessment (i.e., using students' course-based work for outcomes assessment purposes) in the 4000-level literature courses in the Spanish curriculum: the development of a rubric, the use of this rubric in an operational assessment, and an analysis of the results of the assessments and reflections on the process itself.

Assessment Framework

As summarized in Norris and Watanabe (2012), the history of language program evaluation goes back some 40 years, adopting different priorities and procedures in response to overall trends in U.S. higher education. The reemergence of program evaluation activity in the past decade seems to have arisen from a confluence of educational accountability pressures at the national level in elementary

and secondary public education, along with university accreditation mandates that academic programs formulate goals and outcomes for student learning and implement assessments of learning outcomes.

Not all language program evaluation activity in the past decade has been motivated by external mandates, however. In a national survey of department chairs, Watanabe, Norris, and González-Lloret (2009) investigated the evaluation needs, motivations, priorities, and purposes of postsecondary foreign language programs, and results pointed to a diversity of motivations, both external and internal to the respective programs. Although the external pressures for program evaluation were still strong, the results indicated a heightened appreciation for the potential of evaluation to answer questions about several aspects of a program, most notably learning processes and the outcomes of student learning (Watanabe et al., 2009, p. 14).

Whether the impetus for program evaluation activity comes from outside the program or is generated from within, it has repercussions on both the metaphors that underlie the activity and the uses made of the results. Externally mandated evaluations in U.S. education have frequently been cast in an accountability framework—that a college education is analogous to a commercial product; students, and indirectly, members of the public in their function as taxpayers, are analogous to customers whose interests must be protected (Dickeson, 2006). In contrast, internally motivated evaluation is associated with notions of continuous self-improvement that have been captured by Patton's utilization-focused evaluation framework (2008), which he summarizes as follows:

> ***Program evaluation*** is the systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve … program effectiveness, inform decisions about future programming, and/or increase understanding. ***Utilization-focused program evaluation*** is evaluation done for and with specific intended primary users for specific, intended uses. (Patton, 2008, p. 39; emphasis in the original)

As the name suggests, actually using the results of program evaluation activity lies at the heart of the utilization-focused approach. "Doing evaluation" (Norris, 2006, p. 579) or gathering and analyzing data on SLOs, for example, is only preliminary to using the results to make decisions that will improve student learning.

A utilization-focused, continuous-improvement model of program evaluation entails a change of mindset to embrace what Norris (2006) has characterized as a culture of evaluation, which values: learning about how language programs function, identifying the kinds of student learning that result from them, and, most importantly, using those results. Crucial to the process, however, is expertise in the methodology of conducting evaluations or conducting SLOs assessment as a specific form of evaluation. As Norris (2006) observed, relatively few studies on language program evaluation—specifically, on initiatives to assess the outcomes of student learning in foreign language programs—had been published at that time. In response, he and his collaborators have been actively engaged in building the capacity of foreign language educators in this area and providing venues for

publication of the results of subsequent projects (e.g., Norris, Davis, Sinicrope, & Watanabe, 2009; see also http://nflrc.hawaii.edu/evaluation/ for a rich database of resources). This volume is the most recent such initiative.

Project Summary

The outcomes assessment project in the Department of Spanish and Portuguese has, by institutional mandate, focused on the learning outcomes of the undergraduate Spanish major. As stated in institutional documents, the purpose of outcomes assessment at the University of Iowa is "to help maintain program excellence and to prioritize areas for program development" (University of Iowa Outcomes Assessment, n.d.). The results of the assessments in the Department of Spanish and Portuguese will be used by departmental faculty, who are the primary users of the assessment in both their individual roles as instructors of the courses that constitute the Spanish major and their collective role as the body that designs and approves the requirements for the major, establishes departmental policies that affect the major, and oversees programs that contribute to the major (e.g., study abroad programs, the lower-division language program). Ultimately, every student who takes courses in the department will benefit from programmatic and pedagogical changes undertaken as a result of the assessment.

Developing the Rubric

As explained earlier, this case study focuses on the learning outcome for content knowledge of graduating Spanish majors at the University of Iowa, defined as the assessment of students' ability to "analyze, synthesize, and effectively present written information and argumentation in Spanish" (Department of Spanish & Portuguese, 2007). The learning outcome is assumed to develop through course work in literature, culture, linguistics, and film, the four areas in which students take courses for their major. The outcomes assessment team began with the area of literary interpretation and analysis, for which the learning outcome would be demonstrated in "extended interpretive and analytical essays on Hispanic literary texts, using scholarly sources to support their arguments" (Department of Spanish & Portuguese, 2007). The assessment would take the form of an embedded assessment, based on essays students wrote in their most advanced courses.

In the paragraphs that follow, we detail the steps in the process of developing a rubric for the embedded assessment during the 2012–2013 academic year: conducting focus groups with faculty, drafting the rubric, revising the rubric through pilot testing, and finally, conducting the operational assessment by using the rubric to assess essays produced in three literature courses.

Focus Groups

Implementing a utilization-focused outcomes assessment entailed identifying the stakeholders and engaging their active participation (Patton, 2008). We identified the primary stakeholders for the first stage of the outcomes assessment project

as the faculty members who regularly teach upper-level literature courses for Spanish majors. These faculty members would also eventually be the primary users of the assessment and therefore would have important insights into the features of performance that would ultimately be incorporated into the rubric. Ultimately, students—who are primary stakeholders in the assessment project overall—will also be asked to weigh in on the rubric and contribute to their ongoing refinement. However, at this first stage in the process, in which we were constructing the rubric, faculty input was judged to be the most crucial.

Of the nine faculty stakeholders we identified, seven attended one of two focus groups held during the fall 2012 semester, each lasting one hour. At each of the sessions we reviewed the purpose of the meeting and went over the SLOs (see Appendix 1), focusing on the one related to content knowledge. Both focus groups were semi-structured in nature and were based on the same questions and prompts (see Appendix 2). Our objective was to have faculty members talk about their understanding of analysis, synthesis, interpretation, written argumentation, and the use of scholarly sources to support arguments in the context of teaching these 4000-level courses. We also asked the participants to describe their assignments for the major essays they had recently used in in their upper-level literature courses. Following that discussion, we asked them to articulate the features of literary interpretation and analysis in their students' papers that they would consider very good, good, fair, and poor. Finally, we initiated a discussion about the outcomes statement itself, given that it had not been discussed in a group since its approval by the department five years earlier. Ultimately, the assessment team recommended to the department faculty that the outcomes statement be revised to better match faculty practices in and goals for the 4000-level courses in the Spanish major (see following discussion).

Drafting the Rubric

Following the two focus groups, we conducted a comprehensive review of the notes we had taken. We began by sorting the comments into the categories of analysis, interpretation, synthesis, and written argumentation, which would become the elements of the rubric. From these groupings we gained a better understanding of the points of consensus within the focus group participants. Following iterative review of the notes, we organized the comments further to create three levels of performance. At this stage of analyzing our notes and synthesizing them into a rubric, it was important to keep in mind that the rubric would need to be broad enough to assess Spanish majors' abilities in literary interpretation and analysis, regardless of the specific course content.

This process resulted in the creation of a rubric with three performance levels: exceeds expectations, meets expectations, and below expectations (see Appendix 3 for draft rubric). These three categories were more clearly defined than the four categories we had initially discussed in the focus groups. Each of the three categories described the corresponding level of students' analysis, interpretation, synthesis, and written argumentation in a major paper for an upper-level literature course required for the Spanish major.

Meets expectations

Based on faculty comments in the focus groups, we constructed the performance level *meets expectations* to signify that students *analyze* literary texts by dividing them into components, recognizing that the components are units of meaning. Students also use genre-specific textual elements, such as conflict, character, symbolism, and image in their analysis. Students meeting the expectations of the *interpretation* category utilize textual components from their analysis to explain what the text means. To *synthesize* at a level that meets expectations, students engage in dialogue with the primary text and the secondary sources. Students' *written argumentation* at this level is coherent, with a clear beginning, middle, and end supported by textual references.

Below expectations

The rubric draft categorized student writing that was below expectations for Spanish majors as follows: *Analysis* is based on plot summary only; *interpretation* consists of opinions about textual meaning that are not supported by references to the text or to secondary sources; *synthesis* provides little (or no) use of evidence from textual and/or secondary sources, and the writer builds an argument that lacks coherence and/or does not go beyond ideas discussed in class. Furthermore, *written argumentation* that falls below expectations may contain linguistic errors that impede comprehension.

Exceeds expectations

Student writing that falls into the *exceeds expectations* category, according to the draft rubric, includes an *analysis* that divides the text into components, identifies these components as units of meaning, and relates them to each other and to the text as a whole. *Interpretation* exceeding expectations builds a coherent argument for meaning that is grounded in both textual and extra-textual elements. Exceeding expectations in *synthesis* includes recognizing and using a variety of sources as evidence when building an argument. Finally, *written argumentation* that exceeds expectations for undergraduate Spanish majors is persuasive and shows personal engagement with the texts as well as with the topic.

After drafting this rubric, we sent it to the focus group participants via email with a request for feedback. We received only one response; however, from informal conversations we concluded that we had the participants' approval to move forward.

Revising the Rubric

Following informal approval of the draft rubric, the next step was to pilot the rubric with a set of student essays from a 4000-level literature course in the fall 2012 semester. Of the essays written by the 15 students enrolled in the course, 6 essays fulfilled our three criteria for inclusion: (a) the writer had given informed consent for us to read his or her essay, (b) the writer was a declared undergraduate Spanish major, and (c) the essay was an analysis and interpretation of a literary text. We also had a copy of the information about the essay assignment that the course instructor had given to the students.

Sitting together, we examined four of the six papers closely, discussing their features in light of the draft rubric and making notes on the alignment of those features with the descriptors. Through the comparison of the essays to the statements in the draft rubric, it became evident that the descriptors were in need of modification. As we found features of the essays that related to analysis, interpretation, synthesis, and argumentation that were not adequately represented in the descriptors, we adjusted the descriptors by modifying them or adding new ones. The changes to the descriptors prompted us to return to the essays we had already read to review them again.

The process of rating the essays took place in a similar iterative fashion. Our judgments of the essays were holistic (see, e.g., Williamson, 1993), grounded in the discussion in the faculty focus groups, and also analytical (see, e.g., Knoch, 2009), based on a comparison of features of the essays with the descriptors in the draft rubric. Our goal was for our holistic judgment of the overall quality of an essay to correspond to the results of an analytical comparison of the features of the essay with the descriptors in the draft rubric. Altogether, we devoted approximately eight hours to the iterative review and discussion of the essays.

The revised version of the rubric is displayed in Table 4-2. The changes to the rubric that resulted from the process described here are indicated with bold type (additions) and strikethroughs (deleted or moved). The changes to the descriptors and the reasons for each change are explained in the following sections. Given that we had revised the rubric based on papers from only one 4000-level literature course, we were open to the possibility that additional changes might still be needed when the rubric was used for the subsequent assessment. The rubric in its final form can be found in Appendix 4.

Analysis

Writers of essays that exceeded expectations related the text and its component parts to each other and to social and historical contexts. Based on this observation, we added "and to social and historical contexts" to the exceeds expectations performance level.

Interpretation

We found reliance either on plot summary or on quoted material from the text, but not both, in the papers in the meets expectations performance level. However, in the essays that exceeded expectations, writers used both of these strategies to build an effective argument. Based on this observation, we added "textual elements are both macro (plot) and micro (quotes)" to the interpretation category of exceeds expectations.

Synthesis

Some of the writers organized their arguments by characters or events, rather than by themes. This observation led us to reconsider the role of primary and secondary sources in essays at the meets expectations and exceeds expectations performance levels. The wording "uses the primary text and secondary sources

Table 4-2 Changes to the Draft Rubric to Create the Rubric for the Spring 2013 Assessment

	Exceeds expectations	**Meets expectations**	**Below expectations**
Analysis	• Divides text into components, identifies these as units of meaning, and relates the components to each other and to the text as a whole **and to social and historical contexts**	• Divides the text into components, recognizes that these parts are units of meaning • Makes use of genre-specific textual elements (e.g., conflict, character, symbolism, image)	• Plot summary only
Interpretation	• Builds a coherent argument for meaning that is grounded in both textual and extra-textual elements • **Textual elements are both macro (plot) and micro (quotes)**	• Utilizes textual components (from the analysis) to present what the text means	• Presents opinions about textual meaning that are not supported by references to the text or to secondary sources
Synthesis	• ~~Recognizes and uses a variety of sources as evidence in building an argument~~ • **Engages in dialogue with primary text and secondary sources as evidence for building an argument** • **Develops arguments by showing intertextual parallels or contrasts** • **Engagement with sources contributes significantly to the argument**	• ~~Engages in dialogue with the primary text and the secondary sources~~ • **Uses the primary text and secondary sources as evidence in building an argument, albeit with minimal development** • **Repeats central arguments via examples; limited development of the argument**	• Provides little (or no) use of evidence from different sources (textual and/or secondary sources) to build an argument • **Excessive reliance on secondary sources; summarizes sources rather than uses them to produce an argument**
Written argumentation	• Persuasive argument that shows personal engagement with the texts and the topic	• Coherent argument with a clear beginning, middle, and end supported by textual references • **Personal engagement with topic, although only minimally with the text (I agree…, I think…)** • **Essay organized by examples, not by concepts**	• Lacks coherence or arguments ~~beyond those discussed in~~; **thesis is lacking or is not connected to the body of the essay** • Linguistic errors impede comprehension

Bold = added to the draft rubric
Strikethrough = deleted (or relocated) from the draft rubric

as evidence in building an argument, albeit with minimal development" fit best in the meets expectations category, and "engages in dialogue with primary text and secondary sources as evidence for building an argument" characterized the strongest essays and was therefore assigned to the exceeds expectations category.

Once we had moved the "engages in dialogue with primary text and secondary sources as evidence for building an argument" descriptor to the exceeds expectations performance level, we returned to our notes from the focus groups to make sure that this decision was in line with the expectations articulated by the participants. The difference between the meets expectations and exceeds expectations performance levels for this feature of the synthesis category turned out to be one of quality, rather than of presence versus absence. We then returned to the essays in search of evidence of the writers' engagement with the primary text and secondary sources in building their arguments. The result was, for meets expectations, the qualification of the feature with the phrase "albeit with minimal development." We also added a new descriptor to this category, "repeats central arguments via examples; limited development of the argument" to represent a pattern that we were finding in the essays.

Our iterative review of the notes from the focus groups revealed another point in the discussion that had not yet made its way into the rubric: the features of engagement with the text that characterized the best papers from students in 4000-level literature courses. We once again returned to the essays, this time working with the strongest ones, and focused on the writers' synthesis strategies. We did find features that aligned with the comments of the focus group participants, which resulted in the addition of two descriptors to the exceeds expectations performance level: "develops arguments by showing intertextual parallels or contrasts" and "engagement with sources contributes significantly to the argument."

We also found essays in which the writers had made ineffective use of primary texts and secondary sources in constructing their arguments. In some cases they provided relevant examples and quotations, but they did not take the additional step of incorporating the examples or quotations into their arguments. In extreme cases, the writers relied so heavily on a secondary source that they abandoned their own argumentative thread in favor of paraphrasing someone else's work, in one case over several pages. Cases such as these prompted us to add the descriptor "excessive reliance on secondary sources; summarizes sources rather than uses them to produce an argument" to the synthesis category of the below expectations performance level.

Written argumentation

Similarly, in the written argumentation category at the meets expectations level, it became evident that the descriptors in the draft rubric were not sufficiently differentiating between the meets expectations and exceeds expectations performance levels. To make them so, we added two new descriptors to the meets expectations performance level: "personal engagement with the topic, although only minimally with the text" and "essay organized by examples, not by concepts."

In written argumentation at the below expectations performance level, we added a new descriptor: "thesis is lacking or is not connected to the body of the essay," because we saw examples in which the writer began with a clearly stated thesis but then developed an argument that was not connected to the thesis.

To conclude this section, we offer the observation that the rubric we developed for the assessment of the SLO in content knowledge in Hispanic literatures is manifestly different from the criteria that course instructors would use in evaluating essays in their 4000-level courses. Outcomes assessment takes a general, external approach, whereas course-based summative assessments tend to drill down in greater detail to the content and specific skills that have been treated in the course. Both forms of assessment address goals of the faculty involved in teaching Hispanic literature courses, albeit from different perspectives. In their roles as course instructors, faculty assess their students' essays in light of their goals for a specific course, whereas in the context of this assessment project, their role will be to make use of the assessment results in considering areas for improvement in the Hispanic literature curriculum more broadly, beyond the scope of their own courses.

Operational Assessment

With the finalized rubric in hand, we conducted the first operational assessment with essays written in three of the 4000-level Hispanic literature courses offered in the 2012–2013 academic year. As described earlier, these are the most advanced literature courses in the Spanish major, and students must take three courses at this level (although they may choose among 4000-level course offerings in literature, culture, and linguistics) in completing their major requirements.

We conducted the operational assessment using procedures similar to those we had employed in the pilot phase of the project. We began by examining the essay assignment instructions in light of the subject matter of the respective courses, and we asked the course instructors for clarification as needed. We noted that the essay assignments were quite open ended; students were free to select their topics and approach them as they wished. No one directed students to particular texts or even to literary analysis; in all of the courses, some students wrote analyses of cultural or historical phenomena rather than focusing on literary texts. We also noted that students received considerable guidance and feedback at several stages of the writing process. In each course the professor provided the students with feedback on their topics and on a draft well before the final essay was due; in one case the professor also gave feedback on an annotated bibliography. We subsequently reviewed the essays to identify their topics, and we eliminated essays that did not belong to the genre of literary analysis since they were outside the scope of this assessment. Following the procedure we had established for the pilot phase, we also eliminated essays whose writers had not given informed consent for their inclusion in the assessment and those written by students who were not undergraduate Spanish majors. The results of the operational assessment are displayed in Table 4-3.

Table 4-3 Operational Assessment

Course	Essays available for assessment	Essays used in assessment (literary topic)	Performance levels
1	22	1	Below expectations: 1
2	9	3	Exceeds expectations: 2
			Below expectations: 1
3	10	8	Exceeds expectations: 3
			Meets expectations: 3
			Below expectations: 2
Total	40	12	Exceeds expectations: 5
			Meets expectations: 3
			Below expectations: 4

From Course 1, of the 22 essays written by consenting Spanish majors, only one was an analysis of a literary text. In Course 2, of the nine essays written by Spanish majors who consented to participate in the assessment project, three were literary analyses, five related to cultural themes relevant to the course topic, and one was neither cultural nor literary in focus. In the 10 essays from Course 3, eight were analyses of literary texts and two dealt with cultural themes. Utilizing the rubric we had created, of the 12 essays included in the operational assessment, three met expectations, four were below expectations, and five exceeded expectations. We rated the essays independently by comparing them to the rubric. We then discussed both our ratings and our rationale for them to incorporate a final check of the adequacy of the fit between the descriptors in the rubric and the features of the essays. Additionally, we examined the rubric again to ensure that the performance levels were sufficiently different from each other to identify meaningful distinctions in quality among the essays.

To conclude this section, we note that although we considered the assessment described here as the first operational assessment of the SLO in literary interpretation and analysis, it may be more appropriate to consider it the final stage in piloting the rubric.[1] The few changes to the rubric that we made while conducting the operational assessment were minor, and therefore, are not documented here. The most unexpected outcome of the process was the discovery that the writing assignments in these three 4000-literature courses corresponded neither to the departmental policy of requiring 15 pages of formal academic writing in courses at this level nor to the SLO for content knowledge in literature that the department faculty had approved some years before. This and other findings are discussed in the following section.

[1] We are grateful to one of the volume editors for bringing this point to our attention.

Discussion

The stage of the utilization-focused outcomes assessment project described in this case study was enlightening in ways both anticipated and unanticipated. Indeed, the project is ongoing. Following the analysis of the results by the outcomes assessment team, the first author, who is a faculty member, gave a short presentation to the departmental faculty on the development of the rubric and the outcomes of the first operational assessment. Comments from the brief discussion following her presentation are incorporated into the topics addressed in this section: (a) implications of the assessment for the curriculum in the Spanish major, (b) the important role of *doing assessment* in the assessment of SLOs, (c) initial reception of the assessment results by the departmental faculty, and (d) implications of this project for LPDs.

Implications of the Operational Assessment for the Spanish Major Curriculum

The development of the rubric and the operational assessment of essays in three 4000-level literature courses revealed several issues in the curriculum of the Spanish major that had been percolating beneath the surface, perhaps considered by individual instructors in devising their course requirements or discussed informally in small groups but not previously put on the table as policy issues. The assessment project provided the forum for an open discussion of topics that were of interest to everyone involved in the 4000-level literature courses in the major. In the focus groups and in the assessment itself, two issues arose that the departmental faculty will want to address: the writing requirement in the 4000-level courses in the major and the blurring of the boundary between literary studies and cultural studies.

The requirement of 15 pages of formal academic writing in 4000-level courses in the Spanish major

The assessment team approached the collection of student papers for the assessment with the expectation that all of the papers written by Spanish majors who had given informed consent would be included in the assessment, but this did not turn out to be the case. As displayed in Table 4.3, of the 40 essays available for the assessment, only 12 (30%) could be included. We had to eliminate 28 essays because they dealt with topics other than the analysis and interpretation of literary texts. Most of these essays dealt with cultural topics, and not all were academic research papers. The instructor of one course, for example, had given students the option of doing a creative project.

This finding corresponded to a major theme of the focus group discussions—the acknowledgment by the participants that the departmental policy that students in the 4000-level courses produce 15 pages of formal academic writing was no longer meeting the needs or interests of their students, the great majority of whom do not intend to study literature at the graduate level.

In addition, we discovered that the faculty teaching the 4000-level courses in Hispanic literature were interpreting the requirement of 15 pages of formal academic writing differently. From the focus group discussions, we learned that some faculty members distributed this requirement over several assignments during the semester. Others offered students options that did not focus on literary analysis, as long as they engaged in some way with the texts and contexts of the course.

There was consensus in the focus groups that the goal of the major assignments was to encourage students to engage with the course content in ways that were meaningful to them and that different topics and types of assignments would make more sense for students with diverse professional interests (e.g., K–12 teaching, communication studies, psychology, international studies). Clearly, one outcome of the assessment is the invitation to the departmental faculty to reconsider the writing requirement in the 4000-level courses for the Spanish major and, in the case of literature courses at this level, the exclusive focus on literary interpretation and analysis.

The separation of literature and culture in the Spanish major curriculum

The second curricular issue that emerged from the assessment—evident also in the fact that most of the 28 essays were excluded from the assessment because they dealt with cultural topics—concerned a blurring of the traditional division between courses in literature and courses in culture and civilization. Courses in the undergraduate Spanish curriculum have long been classified as literature, culture, linguistics, language or, more recently, creative writing. Spanish majors must take one course each in peninsular literature, Spanish American literature, culture (including film), and linguistics. The tripartite grouping of courses (literature, culture, linguistics) was replicated in the SLOs approved by the departmental faculty in 2007 (see Appendix 1); the knowledge and skills of literary interpretation and analysis and those of cultural interpretation and analysis were understood as distinct learning outcomes. Several years later, the literature courses that faculty are actually teaching appear to have enough cultural content to make that separation less tenable than it had been previously. It also calls into question the separation of literature and culture as distinct requirements in the major as well as the broader issue of whether it is possible to tease apart literature and culture (Kramsch, 1998) or even desirable to do so (MLA, 2007, 2009).

This emerging shift in the curriculum of the major—the infusion of increasing amounts of cultural content into literature courses—appears to represent the confluence of two trends. The first, which is widespread across language programs in the United States, is the movement in the curriculum of the undergraduate major to include fewer traditional literature courses and more courses of interest to today's students—courses that blend literature and culture (Barnes-Karol, 2010)—and revised curricula for the major that stress such areas as translation studies, service-learning, or language for specific purposes (Doyle, 2010; Jorge,

2010; Sánchez-López, 2010). The second trend, articulated in two influential reports from the Modern Language Association (MLA) in the past decade (MLA, 2007, 2009), is the trend toward curricular integration, the creation of a "coherent curriculum in which language, culture, and literature are taught as a continuous whole" (2007, p. 4) and in which "translingual and transcultural competence" (2007, p. 3) is the principal goal.

Summary

These two issues in the curriculum of the Spanish major—the various adjustments to and reinterpretations of the academic writing requirement and the integration of literature and culture in courses designated as literature courses—also emerged throughout the process of developing the rubric. Even if we had concurrently undertaken the development of parallel rubrics for essays written about cultural and literary topics, respectively, we still would have had to eliminate from the assessment 23 final projects—in response to an open-ended assignment that included an option for creative projects—were not formal analytical essays.

Without the outcomes assessment project and these focus groups, these discrepancies between the required course elements and actual faculty practices might have taken considerably longer to surface. Throughout the various phases of the project, it has become clear that the faculty in the department need to rethink the structure of the Spanish major.

The Importance of *Doing* Assessment

The previous section dealt with two important outcomes of the assessment that touch on curricular issues. Various aspects of the process of doing the assessment have implications for departmental teaching and assessment practices that go beyond the confines of the assessment itself.

When the outcomes assessment team met to start the planning process, we proposed searching for rubrics that had been used elsewhere in embedded assessments of SLOs in literary interpretation and analysis, given that Spanish majors in most institutions prominently feature courses of this type. Understandably, we expected that it would be more efficient to use a ready-made rubric than to spend time developing one of our own. Despite an exhaustive review of online resources for the assessment of SLOs in foreign language programs (for a list of resources, see http://nflrc.hawaii.edu/evaluation/R_outcomes.htm#7_1), we were unable to locate a rubric that would accurately represent what we wanted to assess. After developing the rubric for this project (see Appendix 4), however, we experienced a renewed appreciation for the process itself. Even if we had found rubrics developed elsewhere that appeared to meet our needs, there was no substitute for the value of engaging stakeholders (i.e., departmental faculty) in taking ownership of an outcomes assessment whose results they would then discuss, interpret, and use for the benefit of their students and improvement of the program.

The outcomes assessment team found that designing and creating the rubric was as valuable as the resulting product, if not more so. The various stages of the

process—the focus groups, iterative review of notes, the work on the rubric, and the evaluation of students' essays—allowed for some important revelations about the coursework and requirements for the Spanish major at the University of Iowa that likely would have taken a significant amount of time to be revealed otherwise. We fully expect that in the next iteration of the rubric, the faculty who teach the 4000-level courses in literature will be actively engaged in this part of the process as well.

Reception of the Results by the Departmental Faculty

As mentioned earlier, the author who is a faculty member gave a presentation to the departmental faculty on the results of the assessment. Three themes emerged from the discussion: an acknowledgment that the assessment process did indeed uncover issues in the curriculum that are in need of attention, a desire for the assessment team to undertake the process for other content areas in the Spanish major, and a decidedly positive reaction to the rubric. The outcome of the discussion was an endorsement of the work of the assessment team and a desire for it to continue.

The interest of the faculty in the rubric that the assessment team had developed for the assessment exceeded our expectations. Several people expressed interest in using it as a teaching tool, distributing it to students, and initiating discussion on how students could use it to improve the organization and content of their major writing projects. The use of rubrics by students for self-assessment and self-efficacy purposes has been well documented in L1 academic writing across a range of fields (e.g., Andrade, Wang, Du, & Akawi, 2009; Covill, 2012).

Implications for Language Program Directors

Although the outcomes assessment initiative in the Department of Spanish and Portuguese and at the University of Iowa more broadly is focused on the outcomes of student learning at the conclusion of the undergraduate major, an outcomes assessment whose results inform curricular change is likely to have repercussions throughout the language program, initiating what has been termed a "culture of evaluation" (Davis, Sinicrope, & Watanabe, 2009, p. 223). Based on our findings so far, LPDs might draw the following implications from the assessment of learning outcomes for Spanish majors in literary interpretation and analysis.

The value of program evaluation

At the most general level, the outcomes assessment project described here created a site for discussion about issues that concerned them individually but had not been formally discussed. These issues now have higher priority in the topics for discussion and action by the departmental curriculum committee and the faculty as a whole. The same process could be extended to the first- and second-year Spanish and Portuguese language programs: formulate SLOs for the first two years of language study, design modes of assessment, conduct the assessment,

and make use of the results to improve the programs. With the experience behind us of assessing SLOs for graduating Spanish majors, the assessment team would be well positioned to initiate the process in a different program within the department.

Curricular integration

Thanks to the two reports produced by the MLA, discussions on and efforts toward the integration of language, literature, and culture in the curricula of language programs have increased markedly. Well-documented initiatives include the integration of culture and language in elementary and intermediate Spanish courses (Barnes-Karol & Broner, 2010) and the multiple literacies curriculum of the Georgetown German Department (Byrnes, Maxim, & Norris, 2010). (For additional examples, see Norris et al., 2009.)

To be maximally useful, such curricular integration should start at the beginning stages of instruction. As articulated in the summary of the Georgetown German curriculum, "Developing Multiple Literacies," an integrated curriculum is "content-oriented from the beginning of instruction. . . . [It] presents an integration of content and language through oral and written textual genres throughout the undergraduate program" (http://german.georgetown.edu/page/1242716500101.html). LPDs have much to draw on from the key elements in this short summary. Beyond the practical matter of incorporating more literacy activities into programs that for the past three decades have focused primarily on the development of students' oral skills, we also see a challenge to the culture of language programs—to move beyond the language–content divide that has long characterized undergraduate language programs and the departments in which they are housed. The challenge to foster a culturally and textually rich orientation in elementary and intermediate language programs extends to the students as well, because they—well trained by us—have come to believe that language skills are the only legitimate content of their elementary and intermediate courses and may resist the move to a multiple literacies orientation. LPDs will play a key role in effecting such changes.

Rubric development

It is safe to say that most, if not all, elementary and intermediate language courses rely on rubrics for the evaluation of major writing and speaking assignments. In multi-section courses, rubrics provide at least the appearance of objectivity in evaluating the work of students in the same section, as well as uniform standards across sections of the same course. If the LPD builds in training sessions in using a rubric to evaluate student work (often done using exemplars to illustrate the performance bands of the rubric), this standardization can be documented.

This process, however valuable, does not address a fundamental issue in the use of rubrics that we discovered while undertaking the assessment presented in this chapter: the value of developing the rubric that is used in a particular

assessment, in addition to using it. In the assessment of literature students' analysis and interpretation skills, we developed a rubric based on the information we got from faculty of relevant courses on what they expect in students' writing, and then fleshed out the descriptors based on our reading of the students' papers. The result was a rubric that was tailored to the expectations of the faculty and to the features of actual student performances. The quality of the resulting rubric, found principally in its relevance to the use for which it was designed, far surpassed that of generic rubrics that capture the basic distinctions between performance levels but not the nuances.

Given the exigencies of time and schedule, the rubrics used in elementary and intermediate language programs fall into the generic category. They provide a partial fit with the features of student writing or speaking performance, but they have serious shortcomings. Although we do not suggest that LPDs take on the time-consuming task of creating different rubrics for all of their graded writing and speaking assignments, we can suggest two limited but valuable alternatives. One is to collaborate with a graduate student in applied linguistics, second language acquisition, or educational psychology who would be interested in working with course instructors on a rubric development project. Another is to hold brainstorming sessions with the instructors of a course immediately following their use of a rubric to evaluate a writing or speaking assignment. They could be asked to take note of the problems with the existing rubric, and the session could then be devoted to improving it. Both of these options would result in rubrics that are better suited to their intended purpose and, in addition, would incorporate elements of the development process that proved to be so beneficial to us in our outcomes assessment work.

Conclusion

When outcomes assessment is undertaken from an internal drive to gather information that will be used by stakeholders to improve their language programs, rather than to comply with external mandates, it turns out to be a complex and lengthy process. As we have shown, the thoughtful approach comes with challenges. From a practical perspective, it cannot be a short-term project; it is a long-term process that will ultimately address numerous aspects of a program. From a theoretical perspective, such recursive cycles of assessment are to be expected in a program whose faculty have adopted a culture of assessment (Norris, 2006). Engaging in the multistage process of assessing needs, engaging stakeholders and maintaining their involvement over a long period of time, designing instruments and procedures, carrying out the assessment, and analyzing and reporting the results cannot be accomplished quickly. As we discovered, by the time the process is complete and the results are ready for the crucial step of utilization, the needs of the program may have shifted, which will necessitate a return to the needs assessment stage. Although the process is extensive, the outcomes are fundamental to the ongoing well-being of a language program.

Appendix 1. Outcomes Assessment for BA in Spanish (May 25, 2007)

Outcome statement	Knowledge/ skill	Assessment	Feedback to department
1. **Ability to sustain spoken communication at a level sufficient for effective communication with Spanish speakers** Students engage in conversations in Spanish on a range of topics, recount events and personal experiences, and produce explanations and descriptions. (Advanced Low/Mid on the ACTFL scale)	Interpersonal speaking	An OPI-like interview with an appropriate percentage of graduating majors will be conducted with faculty who either are OPI trained or have undergone local preparation for this task. (Direct assessment)	Results will inform discussions and proposals of departmental curriculum committee regarding courses at all levels where speaking and listening in Spanish are course goals.
2. **Ability to analyze, synthesize, and effectively present written information and argumentation in Spanish** Students produce analytical essays in Spanish. (Note: Papers will be classified as belonging to the following types:) a. Extended interpretive and analytical essays on Hispanic literary texts, using scholarly sources to support their arguments b. Extended interpretive and analytical essays on Hispanic cultural phenomena, using scholarly sources to support their arguments	Writing, analysis, argumentation Literary interpretation and analysis Cultural interpretation and analysis Linguistic analysis and argumentation Film analysis	Assessment embedded in more than 170 courses. Students select from these courses two papers written in Spanish that represent their best work. (Papers may be drawn from one or two courses.) The outcomes assessment committee examines student papers according to assessment rubrics for discipline-appropriate interpretive/analytical essays developed by the outcomes assessment committee. (Direct assessment)	Results will inform discussions and proposals of departmental curriculum committee regarding courses at all levels where analysis, interpretation, and accuracy in written expression in Spanish are course goals.

c. Essays that summarize and evaluate analyses of linguistic phenomena, demonstrate the ability to construct a linguistic argument, and report research results in a discipline-appropriate format d. Film analyses on Spanish or Latin American cinema		
3. **Ability to reflect upon one's educational experiences, particularly as these apply to cross-cultural and intercultural issues** Students demonstrate awareness of their growth, as a result of their Spanish major, in (a) academic and intellectual knowledge and analytical ability, (b) understanding of diverse cultural perspectives, and (c) Spanish language skills.	Knowledge building and cultural understanding	Students write a reflective essay (in Spanish or in English) in response to a question posed by the outcomes assessment committee about their overall growth in knowledge, skills, and cultural understanding. In writing their essay, students will draw on a reading of the two papers in their portfolio, as well as their recollections of their experience of the Spanish major, including learning inside and outside the classroom. The reflection will be assessed by the outcomes committee according to a rubric developed by the committee.

Appendix 2. Focus Group Protocol

Overview and topic

* Explain purpose of the meeting.
* Review what we have done so far in the outcomes assessment project: future plans survey, majors survey, OPIs of graduating majors, intercultural competence instrument.
* This year's project: the part of the outcomes assessment project that deals with what students learn in their content courses (nearing the end of their coursework). Explain that the focus of outcomes assessment is cumulative learning, not specific content students learn in individual courses.
* Give some background on the assessment of SLOs at the national level: subject of federally funded programs, edited volumes (including upcoming AAUSC volume), conference presentations.
* Explain purpose of the focus group; overview of work over the rest of the year.
* The main goal is to collect information that will be useful to departmental faculty, so we have to start with their understandings of the key terms in the SLO statement for literature. Gather from them information on how they evaluate students' performance in the essays they write in the most advanced courses.
* Our goal is for them to talk freely for an hour (with a bit of guidance from us) and for us to listen.

Guiding questions

1. Here is the wording of the outcome statement:

   Ability to analyze, synthesize, and effectively present written information and argumentation in Spanish. Students produce extended interpretive and analytical essays on Hispanic literary texts, using scholarly sources to support their arguments.

   Can you first talk about your understanding of the following terms?

   * analysis
   * interpretation of literary texts
   * synthesis
   * written argumentation
   * use of scholarly sources to support arguments

2. What are your writing assignments like? What do you ask students to do?
3. Do your assignments tap the abilities in the outcomes statements, or something else/other skills?
4. When you evaluate students' essays, what features of an essay would lead you to consider the essay very good?

What would make an essay good?
What would make an essay fair?
Describe features of an essay that you consider poor.

5. Does the outcomes statement get at the heart of the learning outcome of a Spanish major? Did the department get it right when it approved this wording? Or is there some other way we could/should have expressed it?
6. Is there anything we haven't talked about that you would like to add?
7. Do you have any questions for us?

Appendix 3. Draft Rubric for Literary Essays

	Exceeds expectations	Meets expectations	Below expectations
Analysis	• Divides text into components, identifies these as units of meaning, and relates the components to each other and to the text as a whole	• Divides the text into components, recognize that these parts are units of meaning • Makes use of genre-specific textual elements (e.g., conflict, character, symbolism, images)	• Plot summary only
Interpretation	• Builds a coherent argument for meaning that is grounded in both textual and extra-textual elements	• Utilizes textual components (from the analysis) to present what the text means	• Presents opinions about textual meaning that are not supported by references to text or to secondary sources
Synthesis	• Recognizes and uses a variety of sources as evidence in building an argument	• Engages in dialogue with the primary text and the secondary sources	• Provides little (or no) use of evidence from different sources (textual and/ or secondary sources) to build an argument

Continued

	Exceeds expectations	Meets expectations	Below expectations
Written argumentation	• Persuasive argument that shows personal engagement with the texts and the topic	• Coherent argument with a clear beginning, middle, and end supported by textual references	• Lacks coherence or arguments beyond those discussed in class • Linguistic errors impede comprehension

Appendix 4. Operational Rubric for Literary Essays

	Exceeds expectations	Meets expectations	Below expectations
Analysis	• Divides text into components, identifies these as units of meaning, and relates the components to each other and to the text as a whole and to social and historical contexts	• Divides the text into components, recognizes that these parts are units of meaning • Makes use of genre-specific textual elements (e.g., conflict, character, symbolism, images)	• Plot summary only
Interpretation	• Builds a coherent argument for meaning that is grounded in both textual and extra-textual elements • Textual elements are both macro (plot) and micro (quotations)	• Utilizes textual components (from the analysis) to present what the text means	• Presents opinions about textual meaning that are not supported by references to text or to secondary sources

Synthesis	• Engages in dialogue with a variety of sources as evidence in building an argument • Develops arguments by showing intertextual parallels or contrasts • Engagement with sources contributes significantly to the argument	• Uses the primary text and secondary sources as evidence in building an argument, albeit with minimal development • Repeats central arguments via examples; limited development of the argument	• Provides little (or no) use of evidence from different sources (textual and/ or secondary sources) to build an argument • Excessive reliance on secondary sources; summarizes sources rather than uses them to produce an argument
Written argumentation	• Persuasive argument that shows personal engagement with the texts and the topic	• Coherent argument with a clear beginning, middle, and end supported by textual references • Personal engagement with topic, although only minimally with the text (I agree..., I think...) • Essay organized by examples, not by concepts	• Lacks coherence or arguments; thesis is lacking or is not connected to the body of the essay • Linguistic errors impede comprehension

References

Andrade, H. G., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *Journal of Educational Research, 102,* 287–301.

Barnes-Karol, G. (2010). Reading (literature) in, across, and beyond the undergraduate Spanish curriculum. *Hispania, 93,* 90–95.

Barnes-Karol, G., & Broner, M. A. (2010). Using images as springboards to teach cultural perspectives in light of the ideals of the MLA report. *Foreign Language Annals, 43,* 422–445.

Byrnes, H., Maxim, H. H., & Norris, J. M. (2010). *Realizing advanced L2 writing development in a collegiate curriculum: Curricular design, pedagogy, assessment*. Modern Language Journal, supplement to vol. 94, monograph. Malvern, MA: Wiley-Blackwell.

Center for Applied Linguistics. (1995). *The Spanish speaking test*. Washington, DC: Author.

College students' use of a writing rubric: Effect on quality of writing, self-efficacy, and writing practices. *Journal of Writing Assessment, 5*. Retrieved from http://www.journalofwritingassessment.org/archives.php?issue=1

Davis, J., Sinicrope, C., & Watanabe, Y. (2009). Foreign language utilization-focused evaluation: Current trends, future directions. In J. M. Norris, J. Davis, C. Sinicrope, & Y. Watanabe (Eds.), *Toward useful program evaluation in college foreign language education* (pp. 57–82). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.

Department of Spanish & Portuguese. (2007). *Outcomes assessment for B.A. in Spanish*. Retrieved from https://provost.uiowa.edu/apps/docs/PDFProvider.aspx?file=plan/Spanish.pdf

Dickeson, R. (2006). *The need for accreditation reform*. Washington, DC: The Secretary of Education's Commission on the Future of Higher Education. Retrieved from http://www2.ed.gov/about/bdscomm/list/hiedfuture/reports/dickeson.pdf

Doyle, M. S. (2010). A responsive, integrative Spanish curriculum at UNC Charlotte. *Hispania, 93,* 80–84.

Jorge, E. (2010). Where's the community? *Hispania, 93,* 135–138.

Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing, 26,* 275–304.

Kramsch, C. (1998). *Language and culture*. Oxford, UK: Oxford University Press.

Lisk in-Gasparro, J., & Leonard, K. (2011). *Assessment of learning outcomes: Speaking proficiency and intercultural competence*. ACTFL Annual Meeting, Denver, CO.

Liskin-Gasparro, J., & Leonard, K. (2012). *Intercultural competence: Study abroad and at-home sources of learning*. ACTFL Annual Meeting, Philadelphia, PA.

Modern Language Association. (2007). Foreign languages and higher education: New structures for a changed world. *Profession*, 234–245.

Modern Language Association. (2009). *Report to the Teagle Foundation on the undergraduate major in language and literature*. Retrieved from http://www.mla.org/pdf/2008_mla_whitepaper.pdf

Norris, J. M. (2006). The why (and how) of student learning outcomes assessment in college FL education. *Modern Language Journal, 90,* 576–583.

Norris, J. M., Davis, J., 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., & Watanabe, Y. (2012). Program evaluation. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Malvern, MA: Wiley-Blackwell. doi:10.1002/9781405198431

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Sánchez-Lápez, L. (2010). El español para fines específicos: La proliferación de programas creados para satisfacer las necesidades del siglo XXI. *Hispania, 93,* 85–89.

University of Iowa Outcomes Assessment. (n.d.). *Outcomes assessment—the University of Iowa*. Retrieved from http://www.uiowa.edu/~outcomes/

Watanabe, Y., Norris, J. M., & González-Lloret, M. (2009). Identifying and responding to evaluation needs in college foreign language programs. In J. M. Norris, J. Davis, C. Sinicrope, & Y. Watanabe (Eds.), *Toward useful program evaluation in college foreign language education* (pp. 5–56). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.

Williamson, M. M. (1993). An introduction to holistic scoring: The social, historical, and theoretical context for writing assessment. In M. M. Williamson & B. A. Huot (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations* (pp. 1–43). Cresskill, NJ: Hampton Press.