Assessing Strategic Cultural Competency: Holistic Approaches to Student Learning Through Media

JUDITH HAMMER
Southern Connecticut State University
Department of World Languages and Literatures
501 Crescent Street
New Haven, CT 06515
Email: jhammer@uconnmail.com

JANET SWAFFAR
University of Texas
Department of Germanic Studies
1 University Station C3300
Austin, TX 78712
Email: jswoffar@mail.utexas.edu

The current study investigated the impact of a German television program on changes in 4th-semester German students’ reflections on cultural perceptions over the course of 1 semester. Sixty-nine students at the University of Texas at Austin watched 4 episodes of the popular German television program Lindenstrasse. After viewing, students were asked to reflect in written response papers on cultural features and patterns of behavior and on cultural differences and similarities. The study results suggest that students’ perceptions of another culture can become more sophisticated when being exposed to authentic filmic material and asked to reflect in writing about observed plot features and cultural manifestations. The key to these results is a strategy for assessing not just students’ recall of cultural content, but also their strategic competencies in negotiating cultural difference. Changes in students’ cognitive styles were tracked by a scale that rewards students’ strategic ability to manage cultural content and sociolinguistic content, including the following categories: (a) rhetorical organization; (b) content; (c) comparative point of view; and (d) interpretive substance. The article provides a model for the assessment of cultural competency (MACC), which can be adapted to assess students’ engagement with the culture represented in various media.

THIS ARTICLE LOOKS AT SHIFTS IN FOURTH-semester foreign language (FL) students’ attitudes and perceptions about German culture in conjunction with testing their viewing of four episodes of the German television serial Lindenstrasse. Post viewings, the 69 participants wrote response papers addressing cultural patterns they observed in the television program. This study compares a sampling of randomly selected students’ papers, written after the first and the fourth viewings of different episodes. The goal of this research was to scrutinize possible changes in learners’ conceptual ability to draw inferences about cultural values—as they defined them—from patterns of behavior writers observed and commented on. An assessment model was developed to identify and classify strategies about how students expressed cultural insights when writing about the televised episodes. That model encapsulates components of “cultural insights” that allow a quantification of rater analyses of qualitative differences in students’ observations.

The categories in this assessment model for cultural competency reflect an interdisciplinary body of recent French and English theory that has fleshed out concepts of culture and language based on their anthropological, sociological, and discursive implications. Notably, cultural theorists such as Bhabha (1988, 1990), Foucault (1972), Bourdieu (1991), Geertz (1973), and Fairclough (1989) share the assumption that language and culture are interdependent and implicated with concepts of power, individual agency, and identity. Concomitantly, research in
psycholinguistics, sociolinguistics, and semiotics. In conjunction with the advent of accessible media worldwide, has contributed to the realization that culture, language, and personal or group identities are dynamic, realized in a variety of cognitive processes and social acts that underlie strategies people use to interface with a culture.


Whereas research has been conducted on the recall and knowledge of FL cultural content (Crawford & McLaren, 2003; Herron, Dubrell, Cole, & Corrie, 1999, 2000; Herron & Hanley, 1992; Storme & Derakhshan, 2002) and analyses of sociocultural pragmatics (Halliday, 1977, 2002; Wierzbicka, 1997, 2003, 2006a, 2006b), few measurements and hence few studies exist that investigate patterns in students' initial perceptions about another culture and whether extended exposure to authentic materials representing that culture changes or expands those perceptions (Wright, 2000).

More specifically, despite the centrality of cultural studies espoused in the ACTFL Standards and the MLA report, the profession still lacks information about what kinds of learning actually result from classroom exposure to authentic materials produced by the FL under study, and what manifestations of cognitive processing might be assessed (beyond correctness of linguistic form and address, and of cultural facts like names and dates). As a result, some confusion persists about types of materials and approaches to learning about culture in an FL classroom; confusion that may well inhibit learning.

Investigating this issue with regard to the study of culture, Chavez (2002) examined students' and teachers' expectations about acquiring cultural awareness in the FL classroom. She found that teachers failed to adequately communicate their approaches to cultural information to their students and that, consequently, students lacked certainty about what they should be looking for and learning about. Chavez's findings support the premise that both the learners' and the course objectives have to be taken into account and rendered explicit if effective pedagogical decisions are to be made. In a similar vein, Menard-Warwick (2009) concluded from her research that although teachers often provide opportunities for students to investigate cultural issues, such a student-centered exploration of culture does not necessarily lead to what she calls interculturality.

To address these problems in the study, the authors developed their assessment model in conjunction with guidelines given as assignment directions for a student response paper to be written after viewing a series of episodes from an authentic German-language television serial. The assessment scale paralleled those guidelines in specifying levels of achievement, and the grades correlated with them, to identify and guide possible changes in the writers' perceptions about culture over the course of a college semester.

STUDY DESIGN

The broader project from which this paper is drawn explored four questions. This report focuses on the data gleaned that address one of the study's central research questions and one of its methodological problems:

1. How did viewing four episodes of a television serial about everyday life on a street in Munich, Germany affect the way learners wrote about German culture in their first and fourth response essays of the four that were assigned as homework after viewing each television segment?

2. Could a scale be developed to assess students' cognitive grasp of the FL television serial—to measure qualitative features of individual writers' expressions that provide information about development in their cultural perceptions and inferences?

To investigate this research question and develop a method for transforming such questions into potentially quantifiable data, a study was designed in which four fourth-semester German-language classes watched a German television program presented with the following sequence of tasks based on materials for support of their learning: (a) a preview in the form of a reading assignment: descriptions or pseudo-biographies of characters to be seen in the Lindenstrasse episode to be shown; (b) a 15-minute preview of the episode in class, with a short discussion of

USING A VIDEO CULTURE

Movies and tele used as tools for language culture, verbal, and acc using filmic texts to detail with the semiotic models and input that gives rise receive patterns. At the semiotic of a vation to students to decipher the prepares student classes. In conse

ual language learn singer, 2001; see: a rich cultural envir of authentic c use (altman, 1989 and appeals to c (fox, 1994; loner. Thus far, how unknown are classroom advantages, what materials can be has been meager. To describe the use practice of listening, (hennessy, 1995; 1998; herron, co heron, hanley, & 1999; see: cules, her 1999). Yet overall about an FL cultu

changes in the resulting from re language they are focus of these stu...
characters' biographies and identification of them in an initial scene from each of the three segments of the episode; (c) a 30-minute viewing of the entire German-language episode without subtitles; (d) a homework assignment to write summary analyses of the episode for the next class hour and (e) a follow-up at the next class with a long-class discussion and review of students' questions and perceptions about the episode and its cultural characteristics. This sequence of activities took place in the 4th, 7th, 11th, and 18th weeks of the semester.

USING A VIDEO SERIAL TO TEACH CULTURE

Movies and television programs are frequently used as tools for learning about the second language culture because they illustrate its visual, verbal, and acoustic expression. Consequently, using films and their perceived benefits dovetail with the underlying assumptions of most semiotic models—they provide more redundant input that gives learners a greater chance to perceive patterns. Arey (1998) argued that reading the semiotics of a filmic text is inspiring and motivating to students. He pointed out that practicing to decipher the meaning of images and texts prepares students for upper-division literature classes. In consequence, film provides an "optimal language learning environment" (Kasper & Singer, 2001; see also Swaffar & Vlatten, 1997), a rich cultural environment (Dodds, 1997), a variety of authentic cultural situations and language use (Altman, 1989; Garza, 1991; Wenkland, 1975), and appeals to cognitive and affective spheres (Fox, 1994; Lonergan, 1984; Straub, 2002).

Thus far, however, research investigating whether classroom use of film has pedagogical advantages, what they may be, and how filmic materials can be used to teach social semiotics has been meager. A number of excellent articles describe the use of film in the FL classroom to practice listening, speaking, writing, and reading (Hennesey, 1995; Herron, Cole, York, & Linden, 1998; Herron, Corio, Cole, & Henderson, 1999; Herron, Hakner, & Cole, 1995; Rifkin, 2002; Rose, 1995; Secules, Herron, & Tomasello, 1992; Weyers, 1999). Yet overall, students' analytical thinking about an FL culture and possible expansion of or changes in their attitudes toward the culture resulting from reading or viewing texts in the language they are studying have not been the focus of these studies. The present article will now address some of these dimensions.

PARTICIPANTS

Students

From the 74 enrolled students during the spring semester of 2007, 69 in four sections of fourth-semester German (GER 312L) agreed to participate in the study. All sections were of comparable size (between fifteen and twenty students). They used the same syllabus and met three times a week (Monday, Wednesday, Friday) in 50-minute sessions for 15 weeks. The class sessions all took place in classrooms with identical media consoles that facilitated the screening of the television program throughout the semester. The viewings and discussions of the television series, as well as students' assignments for this study (such as the questionnaires and response papers), were integrated into the fourth-semester syllabus for standard class credit. Participating students' native language was English, and all had either been placed into fourth-semester German through a placement examination or had gone through the regular progression from first- to fourth-semester German at the University of Texas at Austin.5

Instructors

Of the graduate student instructors of these courses (called assistant instructors or AIs at the University of Texas at Austin), two were female and one was male. The two female teachers were German, and German was their native language. The male instructor, although Swiss, spoke German as his native language as well. All instructors had 3 or more years of experience teaching in the lower-division German curriculum of the Department of Germanic Studies at the University of Texas. The instructors' training included a semester-long supervised teaching of German courses as well as classroom visits each semester conducted by the Curriculum Coordinator.

One of the two female instructors taught two fourth-semester sections. The other two teachers taught one section each. All instructors agreed to participate in the study of fourth-semester courses as designed and to have Lindenstrasse episodes incorporated into the fourth-semester curriculum. All instructors agreed that, with the pedagogical format prepared by the first author of this article, the 8 days of the semester and homework assignments devoted to the study Lindenstrasse would be of value for their students. It was agreed that grades for the response papers would be assigned by course instructors, who would correct them for morphosyntactic accuracy, and that, after the semester was over, uncorrected copies of the
essays would be assessed with different criteria by the researchers. Students were aware that results of this subsequent assessment would not affect their final grade for the course.

Data Collection Instruments

_Lindenstrasse_ was chosen as the authentic text for this cultural study because it is a family serial, arguably a distinctly German genre, that first debuted in 1985 and is still a popular series, with new episodes developed as late as 2008. The serial revolves around the everyday experiences of several families living on Lindenstrasse, a downtown street in Munich. The program deals with a variety of current topics that most native speakers of German can be expected to know and have opinions about, such as immigration, HIV, or conflicts with bureaucracy or in the workplace. It short, it represents a variety of everyday problems dealt with by ordinary people in Germany.

The three instructors of the four course sections were given lesson plans for classroom practices, and student participants received guidelines for writing about their cultural perceptions. These guidelines addressed questions raised in the pilot study. Students in that study voiced the same concerns as the subjects in the Chavez study discussed earlier. Despite the fact that study raters' assessment of student essays' cultural content would not affect their course grade, learners wanted to understand the research goals of the assignment and what concepts of culture the raters of these papers sought.

In response to these concerns, semistructured directions for the response papers were employed for the main study. These guidelines asked students to identify main features of the content of the episodes in _Lindenstrasse_ and to identify and explain any cultural features they saw manifested in various reactions of different characters to the events and behaviors depicted (see Appendix A). All these essays were to be written in German, and students were instructed not to use online translators or to receive external help.

The goal of collecting the response papers was twofold. First, it was hoped they would reflect cultural issues students perceived as a result of watching _Lindenstrasse_. Second, papers that contained students' personal reactions to the issues discussed in _Lindenstrasse_ might reveal their initial horizon of expectation about German and U.S. culture, as well as the cultural topics addressed in the televised serial. Thus, a comparison of their initial and last papers might show conceptual and attitudinal changes that students expressed after more extensive viewing. Possibly, students' written responses might reveal whether they were able to observe more precisely, in the course of one semester, plot features of _Lindenstrasse_ and whether they saw and understood the related values, attitudes, and behaviors suggested by the serial's depiction of different ethnic, age, gender, and socioeconomic groups in Germany. To identify the presence or absence of such changes in students' articulation and thinking style, a holistic scoring scale to assess students' response papers was developed by the present authors and Katherine Arens; the scoring scale measurement will be discussed later.

IMPLEMENTATION

Instructions to Teachers and Students

To regularize instructions, administration, and pedagogical uses of _Lindenstrasse_, including all pre-, during, and postviewing discussion activities, instructors were provided with the same lesson plans (see Appendix B). The first author of this study visited all classes to identify any variations or problems that might occur. In the main, the recommended procedures were followed consistently. These plans included activities, instructions about how to conduct those activities, the purpose and goal of the activities, and additional background information for instructors and students.

Pre-Viewing Activities

Students first read short biographies of the people appearing in the respective _Lindenstrasse_ episode as a homework assignment. The subsequent class session started with groups of three or four students who reviewed what they recalled about these characters, prompted by a list of their names on the blackboard. This activity served as an advance organizer by introducing these characters as participants in situations and contexts within an episode. As noted previously, all _Lindenstrasse_ views were in German and without English subtitles. After viewing, students were asked to write response papers to the episode as a homework assignment. The response papers were between 1.5 and 2 pages in length, written in German, and handed in at the outset of the following class session (see Appendix A).

Postviewing Activities, Day 2

On the day after the initial viewing, all classes started with groups of three or four students who talked to each about the cult their response group pooled its categories on topics and wrote notes. The students then the category or religion. Stud discussion group notes on the board were asked to turn their chosen topics and instructor displayed their points of view (see Appendix C).

The Assessment Scale

A scale to preclude cultural competence, whether and diagnosis could also mean an increased awareness, written reflection, or the identification of areas of concern. Appendix C identifies the type of feedback that underlies the whether students are able to observe behaviors in the metric as cultural expectations.

The concepts of culture with reference. The ACTFL's Standard in the 21st century was possible learning, performative Standards define five interlocking comparisons, Con cause the characters to be based on a broad set of practice and the information in patterns and content, these rules of performance, dialectal organization necessary to maintain.

The MACC scale says, formulating accountability (logical relationship...
talked to each other for approximately 5 minutes about the cultural features they mentioned in their response papers. After that discussion, the group pooled its information and created cultural categories on the blackboard, and then used them to sort features from the episodes under appropriate headings. For example, if the cultural feature the students noted was circumcision of Muslims, then the category created was practices of Islam or religion. Students assigned themselves to a new discussion group according to the cultural feature written on the board they wanted to discuss. They were asked to use 15 to 20 minutes to consider their chosen topic, using discussion questions the instructor displayed on the classroom document camera (see Appendix A).

The Assessment Scale

A scale to provide a model for assessment of cultural competency (MACC) was developed to see whether an instrument designed to identify and diagnose students’ perceptual strategies could also measure whether learners exhibited an increased awareness of self and other in their written reflections on the cultural information they identified in the "Lindenstrasse" videos (see Appendix C). More specifically, the typologies that underlie that scale focused on ascertainment whether students were able to increase their capacity to observe possible trends, attitudes, or behaviors in the video that they would characterize as culturally similar or dissimilar to their expectations.

The concepts at the basis of the rubrics define culture with reference to explicit language performance. They were derived with reference to the ACTFL’s Standards for Foreign Language Learning in the 21st Century, understood as a set of possible learning outcomes linked to language-based performance in cultural contexts. The Standards define these contexts as a set of five interlocking domains: communities, cultures, comparisons, communications, connections. Because the class reported on here was defined as language-based and interested in cultural competence and the ability to process cultural information in patterns rather than in any specific content, these rubrics stressed the language base of performance, defined here as aspects of rhetorical organization and the elements of language necessary to manage them.

The MACC scale for assessing differences in essay formulations was divided into four categories of accountability (a) rhetorical organization (the logical relationships among topics and plots of the episode); (b) verifiable content (accurate replication of plot features and concrete information); (c) an identifiable writer point of view (distinguisihing the point of view of the episode characters from that of the writer); and (d) interpretive substance (examples of inferential reasoning, showing how they drew conclusions on the basis of the concrete information).

The scores for each category in the MACC’s grading scale emerged when raters identified and awarded points for the presence, partial presence, or absence of statements or chains of expressions that met the descriptions of the three to six rubrics identified for each of the four categories of accountability on the scale. These categories and the individual rubrics under them served a guide to a holistic evaluation of the writing when they identified the relative presence or absence of the cognitive strategies students used to manage details both of cultural knowledge and of sociolinguistic content in the video episodes.

Points achievable for each rubric ranged from one to three, with one reflecting an absence or lack of clarity in reflecting a rubric, two indicating that a rubric was partially expressed, and three an intelligibly articulated (but not necessarily linguistically correct) expression of a cognitive strategy in the MACC. To establish agreement between raters, comparisons were made on the basis of total scores awarded in the four categories rather than on individual scores awarded for the seventeen individual rubrics, with different kinds of cognitive activity reflected in each. Note that the cited illustrations, taken out of context, have been translated from the German. To compare the language use and strategies of a sample first and fourth essay, see Appendix D.

THE MODEL’S ASSESSMENT CRITERIA

Category 1: Rhetorical Organization

Assessing the linguistic dimension writers use involves measuring the degree to which they employ effective communication strategies. Whether writers can set and hold a position or a point of view about another culture, for example, will depend on the scope and accuracy of their comprehension and their capacity to produce sentences that convey their grasp of the materials heard, read, or viewed. Consequently, rhetorical organization was identified and defined as the coherence and intelligibility of language used to express ideas about culture in a written statement, as an essential rubric or component of any research.
assessment about shifts in students’ cultural horizons.

Thus, the measurement rewarded fundamental components in rhetoric: a dominant topic or argument framing the paper and the presence of subtropics related to that argument. For coherence, transitions and discourse connectors between sentences and paragraphs would also be assessed, as would the ways in which details from the video segment would illustrate the topic addressed. Such features were not taught specifically in conjunction with the Lindenstrasse project, but they were reviewed in course materials during the semester.

The first rubric (Rubric 1) in this category of four strategies defines rhetorical organization and allows raters to award points on the basis of correlations between language and content. The first rubric awards points for simple rhetorical competence: for single or related sentences that present the main subject in each of the three episodes and the extent to which any supporting sentences at the outset of the essay identify the writer’s concept of the global issues and specific themes addressed in episodes of the television segment.

The second rubric (Rubric 2) looks at the reader’s sustained engagement with the concepts identified in Rubric 1, rewarding subsequent reappearance of these themes, major arteries in the essay’s roadmap. In other words, the theme of this “topic sentence” has to reassert itself as a thread in subtopics throughout the essay, showing an attempt to argue a case across more than one sentence or clause boundary. To illustrate, an opening sentence such as “Lindenstrasse is a German soap opera that is very popular” could be awarded two points as a general topic sentence, but no more unless subsequent sentences linked that popularity to specific stories in the episode. If, for example, that statement were expanded in at least one of the following paragraphs by the explicit or implied assumption that the aspects of life depicted and chief features of the episodes’ segments are interesting to Germans, the writer would be awarded three points for Rubric 2, “sections of that essay each related to its major argument.”

The textual form of these assertions in the students’ papers did not need to be elegant or all-inclusive to receive credit, since the rubrics reward content not grammatical correctness. A subsequent paragraph might start with the topic sentence: “Appropriate lots of Germans watch a program that shows some Muslim traditions and a German woman who wants to marry a Muslim man.” This assertion suggests that the topic “popularity” can be scored as developing the theme and potential argument “implications of popularity for the German viewing audience”—“lots of Germans watch…” restates the assertion that such topics are relevant in Germany today.

The rhetorical feature assessed in Rubric 2 is whether any part of the writing introduces subsequent topic sentences in subsequent paragraphs that link the popularity of Lindenstrasse to writers’ views about how German and, possibly, U.S. audiences might see religious differences or the dynamics of traditional and nontraditional relations depicted in the video segment. If the writer digressed midway through the essay or failed to link this initial focus to further segments of the essay, a point would be deducted. The total absence of such thematic threading would yield only one point for this rubric.

The third rubric (Rubric 3) marks, “transitions between phases of argument.” This rubric awards points for discourse markers such as “first we see,” “next,” “a very serious problem emerges when…”—words or phrases that convey to the reader the sequence of topical events or the importance the writer assigns to them.

The fourth and last rubric (Rubric 4) in this measure of rhetorical organization looks at whether details from the source text are presented with referential clarity. In reference to the prospective Muslim wedding, for example, does the writer identify who is marrying whom and what friends and family think about it?

This last subrubric rewards the degree to which “details from source text are subsumed to (sections of) the argument.” Although related in principle to the second category, “verifiable content,” the MACC’s designers focused here on the way essay details lend texture and depth to the main messages of a segment. If the writer’s argument is, for example, that Islamic traditions lead to confrontations between Germans of a non-Islamic background and their Islamic neighbors, the essay needs to elaborate briefly about the reactions of other characters to a woman wearing a headscarf or having her 4-year-old boy circumcised. As with all rubrics in the scale, raters awarded three points for the presence, two points for a partial answer, and one point if such illustrations were absent.

**Category 2: Textually Verifiable Reference to Content**

If culture is to be learned about and understood, observed examples of that culture must be accurately comprehended and represented. Consequently, a rubric for content was developed to measure whether w minor issues, facts, or situations b in comparisons became generalizations or if how sequences, relationships between accurate or inaccurate inferences frequent facts. As a related option was included in the rubric, the present information about a

The category of a degree to which situations and topics are diagrammatically and accurately described and illustrated. Designers were instructed to consider the following aspects of student work: Did students use diagrams to depict the narrative structure of the text? Was the diagram accurately created and labeled to include the main points of the text? Was the diagram visually appealing and easy to understand? Did the diagram effectively represent the key elements of the text, such as character relationships, plot development, and setting?

The rubric was designed to assess the accuracy and completeness of the diagrams. Points were awarded for the inclusion of all necessary elements, the correct representation of the events, and the clarity of the diagram. The rubric was also designed to assess the students’ ability to use the diagram as a tool for communication. Points were awarded for the clarity of the labels, the legibility of the text, and the overall effectiveness of the diagram in conveying the intended message.
measure whether writers recounted major and minor issues, facts, and situations accurately. Accurate distinctions between major and minor issues, facts, or situations become significant first steps in comparisons because they can forestall cultural generalizations or stereotyping. In this same vein, how sequences, chronology, and the relative relationships between events are replicated reflects accurate or inaccurate perceptions of data. Flawed inferences frequently rest on misconception of facts. As a related consideration, genre recognition was included in this rubric because genre influences the presentation and selection of factual information about a culture.

The category of rubrics as a whole assesses the degree to which students comprehended situations and topics depicted in the television program. Designers wanted to look at how fully and accurately students recalled the program and described it. This segment of the scale was weighted twice as heavily as the other categories because content recognition establishes the basis for higher-order thinking about culture. Such thinking involves synthesizing text content and student background information into topics and analyzing them from different points of view. If students fail to give a correct account of the participants and issues presented in Lithendenstrasse, or merely generalize without examining more closely the features of different people, places, events, and problems presented in the episodes, their inferential reasoning about culture suffers.

To assess the basis for pitfalls, raters looked first at the degree to which writers understood and reported an episode’s situations and facts coherently. Did they, for example, note the connection and relation of the overarching and more important events in a segment? Was the topic “circumcision” related to details about the event with indications about who was invited, where the event took place, or the kinds of interactions that occurred between guests?

For such appraisals of content articulation, rating considerations similar to those used in assessing rhetorical organization applied. For Rubric 1, descriptive generalizations without specifications did not suffice. For example, the correct assessment that “[h]is first story is about all the people involved in a ritual” would receive only one point if not supported by subsequent statements (“in this case...”) about the type of ritual or social issues pertaining to this event. In other words, if an observation remained inconclusive or failed to exemplify or amplify the learner’s management of the broader content of the video segment, it did not count for more than one point.

Alternatively, when students identified the segment features that illustrate the rituals previously described they received two points if they defined “ritual” or referred to events (“circumcision”), attitudes expressed (“the young boy was afraid”), or issues raised (“the real father’s rights”). If such additional, more detailed references characterized discussions of at least two of the major issues the student’s essay identified, it would receive three points for Rubric 1, “major issues/facts/situations from the source text accurately recounted.”

In sum, the content reference for Rubric 2 awarded points for describing details of the major content in either a real-time or flashback sequence that replicated the original. Mentioning the fact that the non-Muslim fiancée must have her stepmother present when she meets her prospective Muslim in-laws, or the fact that she prepares an extensive dinner in her own kitchen, speaks to the formality and family significance of the betrothal—a clear specification of the basic situation “ritual” that elaborates about a “Muslim ritual in Germany.” Not all, but at least two or three, of such details had to be found in the students’ description of a particular episode from the video in order to garner three points.

Content reference Rubric 3 asked raters to assess how the student essay weighted such major and minor issues—whether the writer recognized the difference between super- and subordinate features. Three-point ratings did not depend on comprehensive attention to all factors in the text, simply to evidence that the writer linked details to overarching issues. For example, the fact that the fiancée was so preoccupied with cooking dinner that her young son felt neglected speaks to the problem of new priorities introduced when new traditions are adopted. The essay that consistently linked such details about the son’s dilemma to the major topic “two different traditions” received three points for this rubric. Conversely, an exegesis about parental neglect that digressed from the content of the episode would receive no more than two points, such being an issue imposed on the text, that is, extraneous, although not completely alien to it.

Content Rubric 4 addressed the essays’ use of events, nouns, adjectives, and verbs as semantic representations of particular features—whether the writers’ choices avoided distorting or stereotyping events and people. As with all rubrics for this level, the raters penalized inaccuracy. Thus, for example, a reference to an unmarried mother as “immoral,” with two or three other such judgmental adjectives occurring in reference to other people or events, would cost at least one point if
such judgments failed to represent statements in the video segment.

Content Rubric 5 is distinguishable from Rubric 4, an explicit judgment, when writers made imprecise or inaccurate inferences. Assertions such as “everyone dislikes this girl” or “the money was stolen” were penalized by one point if only one person disliked the girl in question or thought that a sum freely given had been obtained illegally.

Rubric 6, the last piece of the content category, rewarded observations about how the genre (whether considered as a soap opera or a family serial) influences the way content is presented in Lindenstrasse. Only two points were awarded if students simply mentioned that the television program was, for example, a soap opera (“Soap operas are interesting but they aren’t correct [i.e., accurate representations of society]”). For three points, students were expected to exemplify an assertion, by considering the impact of the genre with elaborations such as “viewers expect more dramatic events and exaggerated behaviors.” Such an elaboration requires an indication of how genre affects what is observable and suggests that the writer interrogated the episode’s content in terms of its genre.

Category 3: Comparative Point of View

The third set of rubrics focused on whether student writers recognized similarities, but also could distinguish between features of the C2, or target culture, and those of the C1. In other words, did writers make comparisons or contrasts with their own culture when reflecting on features of the German culture they saw represented in Lindenstrasse. Here, recall and discussion of particular information is not at issue; instead, raters rewarded the writers’ ability to identify behaviors, values, and interactive patterns in the video segment that struck them as anticipated, new, or different from their expectations, or how writers assessed what they saw as stereotypical.

The four rubrics in this category rewarded observations students made about their own culture, the target culture, or vis-à-vis the target culture. The first and second subrubrics, “source and target cultures clearly differentiated where appropriate” and “clearly equated where appropriate,” assessed whether, on the basis of their cultural origins, writers attributed to themselves similar or different expectations than those made when they commented on events and behaviors in Lindenstrasse. Thus, when writers noted that the apartments depicted are relatively small, did they also observe overtly that U.S. apartments for comparable socioeconomic groups tend to be somewhat larger?

In contrast to the content assessment, here the scoring did not depend solely on the accuracy of the observation, but on whether the writer identified a particular comparability or contrast between both cultures in terms of economic, religious, or social factors about which the writer made explicit references. If several such observations occurred in an essay, the statements would be awarded three points if factually correct; two points if the raters found them inaccurate or flawed—if, for example, the writer assumed that most Turks are Christians. In other words, as long as the writer expressed a point of view and awareness of a cultural issue, partial credit was granted.

The reasoning behind allowing partial credit for factually flawed comparisons is that learning to draw inferences about cultures is a process of learning how assumptions underlying inferences affect subsequent interpretations and conclusions. For the scale to serve as a feedback or diagnostic process for teachers and students, raters found it essential to reward in part what may otherwise be considered a factually flawed assertion (e.g., most Turks are Christians) so as to encourage the students’ effort to make an inference, but at the same time point out potential pitfalls for observers of another culture—the difference between informed assumptions about the FL culture and flawed inferences that are based on inadequate information about another culture’s way of dealing with ostensibly similar issues. If assumptions in an essay are based on cultural facts not found overtly in the textual material students work with, the error is attributable to student background rather than course content. Accuracy of course or text content has been assessed in the category “referential accuracy”; these point-of-view rubrics diagnose indices about awareness of one’s own point of view versus that of others, perspectives that constitute the student’s basis for comparing two cultures.

Point of view Rubrics 3 and 4 look at how writers actively “interrogate stereotypes,” how and whether they position themselves either as outsiders within their own culture (Rubric 3) or as a surrogate insider observing the foreign culture (Rubric 4). Given the brevity of these essays—on average between 600 and 750 words—both raters agreed to grant two points for one observation and three points for two or more. Thus, two or more outsider comments such as “With their public transportation systems, I guess older Germans can get to places easier than older

Judith Hammer on Americans” or “I1 as Christians, but people who have lived as evidence of one’s own culture

Similarly, subtle criticism in term “source text unclear”) might receive a German depiction of a Muslim
between Rubrics 3: judgmental behavior episode represent
ation of stereotypical viewpoint.

Category 4: Interpretation

The last rubric on ideas and comments cognitively the m effectively across transcultural negotiation process Other where apper
voices knowledge representations of behavior for class or into where two co
parallel structures one’s own culture, an FL text reflects

For sophisticated able to interrogate well. Significantly, a how a single inst
broader pattern a this point that writ
scultural pheno

This category le or implications wr
analyses—conclu
content and
they saw as signific
three points needed to pr
video episode. Th
the knowledge beyond the literal argument, proba
ment with concre

Rubric 1, “drawing of parallels,“
Americans” or “I have always thought of Germans as Christians, but Lindenstrasse definitely shows people who have different religions” were credited as evidence of awareness about stereotypes in one’s own culture.

Similarly, students who expressed concerns or criticism in terms of the episode (Rubric 4: “source text clearly positioned within source culture”) might receive points for concerns about a German depicted in the serial who denigrated a Muslim neighbor. The difference between Rubrics 3 and 4, then, is that identifying judgmental behaviors or views depicted in the episode represents a different locus for interrogation of stereotypes than does an overtly U.S. viewpoint.

**Category 4: Interpretative Substance**

The last rubric draws on features of the comparisons and communities Standards, the two that are cognitively the most complex. To communicate effectively across communities, a writer negotiates transcultural understanding responsibly. That negotiation process reflects awareness of Self and Other where appropriate, an awareness that involves knowledge of such complexities as accurate representations of differences in language and behavior for class or gender. It also reflects insight into where two cultures can be viewed as having parallel structures, awareness of stereotypes from one’s own culture, and whether information from an FL text reflects stereotypes.

For sophisticated comparisons, writers must be able to interrogate clichés in their own thinking as well. Significantly, such writers hypothesize about how a single instance may or may not reflect a broader pattern of behavior or attitudes. It is at this point that writers are poised to interpret transcultural phenomena and their implications.

This category looks at what broader inferences or implications writers saw as resulting from their analyses—conclusions based on accurate replication of content and the visual and verbal behaviors they saw as significant in the television episode. To gain three points in any of these rubrics, writers needed to provide the basis for their conclusions by referring accurately and with some precision to behaviors and attitudes they identified in the video episode. This category thus points to synthetic knowledge—to the writer’s ability to move beyond the literal or mechanical and manage an argument, probably by substantiating the argument with concrete data.

Rubric 1, “draws conclusions beyond description of parallels,” rewards substantiated inferences and penalizes unfounded generalizations. Thus, concluding statements such as “[the behavior of people in the show is typically German... mostly conservative” lost a point as an unsubstantiated generalization, whereas the comment “[In Lindenstrasse,” many Germans seemed to accept Turks as members of German society, inviting them to social events and intermarrying” qualified as the writer’s substantiated conclusion.

Rubric 2, “interrogates clichés/stereotypes in source text,” looks for statements that indicate writers are reflecting about what they see as the prejudices or too-facile generalizations exhibited in the video episode itself. This rubric proved to be the least frequent one that raters noted in the essays. One of the exceptions was the statement “Muslim religion plays an important role in this episode, but the people concerned didn’t talk about the religions ideas, just circumcision. Islam is complex and this sequence treated the religion in black and white terms.” This writer positioned her comment as an impartial critical viewer rather than formulated as a criticism as a U.S. viewer.

Rubric 3, “interrogates clichés/stereotypes in own thinking,” had more frequent exemplification in the papers scored. Writers mentioned their surprise about seeing Germany as a multicultural nation. Some criticized the United States for its lack of readily available public transportation systems, while others were impressed with the personal interest a doctor in Lindenstrasse took in his patient and reconsidered their prejudice against socialized medicine. Where only an implied insight was inferable, for example, praise for the doctor, but no overt connection to comparative medical practices, the writer received two points. When the observation resulted in an explicit confrontation with the writer’s presumptions, three points were awarded.

**RESPONSE PAPER RATINGS**

To establish whether the foregoing scale could recover and in some measure quantify the qualitative differences among student response papers, sampling was done from those essays written after the initial viewing of Lindenstrasse in the fourth class week and those written by these same students after the final viewing in the 12th week of class. Two raters scored a group of essays independently and then worked together to establish a mutual understanding of how to use the criteria to rank differences from one to three. After about 2 hours spent discussing the basis for variation in assessment of the four categories described previously, raters were able to undertake subsequent
grading independently with no more than a 10% difference in rankings, indicating a 90% independent agreement in applying the scale.

In evaluating the students' statements, raters discovered that although they had weighted individual categories somewhat differently, their cumulative scores under the four categories proved to be within one point of difference. Consequently, it was only when the total scores in categories were not within 95% of each other that raters addressed variations in rubric assessment within categories by comparing their readings of how an essay dealt with individual rubric features.

**TABLE 1**

<table>
<thead>
<tr>
<th>Category</th>
<th>Essay 1</th>
<th>Essay 4</th>
<th>Possible Points on Scale × 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rhetorical</td>
<td>114</td>
<td>149</td>
<td>192</td>
</tr>
<tr>
<td>Organization</td>
<td>184</td>
<td>237</td>
<td>288</td>
</tr>
<tr>
<td>Content</td>
<td>110</td>
<td>140</td>
<td>192</td>
</tr>
<tr>
<td>References</td>
<td>105</td>
<td>124</td>
<td>192</td>
</tr>
</tbody>
</table>

Quantitative Analysis of Response Papers

Thirty-two response papers written in the 4th and 15th week of class were analyzed. To make this selection, two women and two men from each class (a total of sixteen students) were randomly chosen from the sixty-nine students in the four classes that participated in this study. The MACC scale assessments indicated that these students' written responses improved between the 4th and 15th weeks of the semester. That improvement occurred to varying degrees with different writers and with respect to different rubrics.

Together, the sixteen students improved most in the categories “rhetorical organization” (35 points or an average gain of 31%), “textually verifiable reference to content” category (with a 53-point or 28.7% average gain), and “point of view/comparative cultural competence” (with a 30-point or 27.5% average gain). The notably smaller average gain was in “interpretive substance,” which increased by only 19 points, an 18.2% average gain. Overall, students' average score on all categories increased by 26.5%, reflecting a total of 137 points. Table 1 provides totals of the sixteen students' scores on their first and fourth response papers with respect to the scale's four categories of rubrics. See Appendix E for individual scores in each category, with total scores and averages noted in descending order.

Some caveats with regard to these results must be noted: The MACC scoring attempted to provide a holistic assessment of culture's multiple cognitive and linguistic domains under a single umbrella. As such, the domains and the weighting of those domains are subject to caveats inherent in any effort to put theory into pragmatic practice. In discussing the findings, no claims can be made for more than inferential findings, since the authors' measurement instrument, and results, would first have to be subjected to replication studies. That said, the assessment differences in these preliminary data suggest that changes in the cultural horizons occurred in the majority of students whose papers were rated. That finding raises the possibility that this progress may be measurable with a comprehensive cultural assessment model.

In this context, it should be emphasized that progress was not consistent for all students across all categories. Although all but one student made progress in at least one of the Categories 1 and 2, degrees of gain varied from one to eight points. The individual score comparisons in descending order found in Appendix E reveal to what degree the score gains and losses in different categories varied among individual students. Two writers, Andrew and Corey, had the most consistent gains in all categories. In Category 3, Point of View, five students' scores were flat or slightly lower by a range of one to three points.

Summary

This study also revealed that a majority of students be broadened due when watching an improvement in MACC may have. All essays' grade point areas suggests that students are relative difficult, primarily managed reproduction of. These cognitive are a point of reasoning (see Anderson & McKechnie, 1995).
improvement over the course of a semester, the MACC may have diagnostic value.

All essays' greater success in the categories of rhetoric and content than in assuming a particular point of view or drawing cultural inferences is not surprising, given what is known about relative difficulty in cognitive processing. As previously noted, perceptions of content are cognitively managed through recognition and written reproduction of what has been comprehended. These cognitive processes are generally considered by psycholinguists to be less demanding than the analytic and synthetic reasoning necessary to assume a point of view or to engage in inferential reasoning (e.g., as encompassed by Bloom's Taxonomy [1956] and its subsequent modifications [see Anderson & Krathwohl, 2001; Marzano & Kendall, 2007]). True, reproducing recognition in a FL requires cognitive synthesis from writers, but as the video story was introduced and heard in German, translation would itself be a factor in essays, whether written in German or English. The authors also assume that the progress students made during the course of the semester doubtless contributed to improved competencies in their expanded cultural perceptions expressed in later viewing of Lindenstrasse. By the same token, the four repetitions of the viewing and postviewing procedures also presumably enhanced later, higher-than-average performance as measured by the MACC.

At the same time, because the scale is designed to assess learners' reasoning and conceptual management, it rewards abilities less likely to be measured when a FL course focuses more on language acquisition in and of itself. On the MACC scale, a level of language ability is the basis for comprehension and oral or written performance and, as such, it mattered, but probably accurate usage and word choice weighed less predominantly than is usually the case in the intermediate-level FL class. Consequently, this research raises the question: Should college-level FL programs reward students for being able to communicate insights about what they learn from their exposure to FL materials, apart from their linguistic abilities per se?

SUMMARY

This study allows for several tentative findings. First, it revealed that in the raters' sample, a majority of students' cultural horizons appeared to be broadened during the course of one semester when watching and discussing Lindenstrasse. Most learners expanded their reading and articulating of images and cultural content they viewed in the televised text. For most, repeated exposure to the video led to increased awareness and ability to articulate a variety of cultural similarities and differences identified in the model for assessing cultural competencies.

The study's second finding indicates that it may be possible to look at qualitative differences in learner perceptions about culture by quantifying strategies (discourse, social, cognitive) writers used to identify cultural phenomena. That said, a number of important qualifications must be made with regard to these findings and the model outlined here.

First, other types and selections of materials and different presentational trajectories might reveal different patterns of progress or, indeed, failure to gain additional insights into manifestations of culture. Second, the model itself will have to be studied and replicated under similar conditions to see if it can be verified and be considered a reliable tool for assessing students' perceptions of texts produced in different formats and genres—specific problems in different texts may require different specialized rubrics under the general categories. The authors' assumption that materials produced for native speakers of the FL studied are preferable to those written for pedagogical purposes might also be tested. In short, although used here as an instrument for assessing student progress over the course of one semester, the model has possible applications and variants for longitudinal studies and use among different FL students at different learning levels and institutions.

The work presented is drawn from the first author's larger study; additional data analysis and the limitations of those results are discussed in Hammer (2008). In this article, the authors focused on the cognitive processes revealed in essays in which students were asked to make inferences about social systems represented in a family video series. The process employed for identifying the various strategies used, the MACC scale, was designed to pinpoint the discrete discursive, social, and cognitive abilities of postsecondary students. These findings cannot reflect the cognitive work of children or secondary school children. Preadolescents and adolescents would not be in a position to utilize several of the strategies found in essays produced by fourth-semester college students or at least not to the extent demonstrated in the essays discussed.

These caveats impinge on both the pedagogical and assessment implications of the rubric-based model presented here. If employed in further
research studies or pedagogical implementations, a variety of modifications would doubtless occur. It should be emphasized, then, that regardless of future applications of the model described in this article, the premise behind its construction and application in this article has been to open a wider discussion about assessments of FL students' perceptions about culture and other content areas. Assessment for this realm has become essential if the profession is to responsibly account for these students' advanced foreign language capacities (Norris, 2006, p. 167) and their translilingual and transcultural competence (MLA Report, 2007, pp. 5-4). In other words, in this era of transition in the status and praxis of FL programs in the United States, FL study's critical need is to assess student performance in those domains the profession sees as its raison d'être and, at present, cultural studies have been situated at the heart of that enterprise.

NOTES
1. For a full account of the study and its ramifications, see Hammer (2008). 
2. This question, #1 in the original study, asks: Would students first and last response essays reveal differences in the way students organize their ideas, recalled content, identified their writer position on events viewed, or identified more of their implications? The design’s second question asked about the effects of the use of English or German in discussion; the third question inquired about differences in classroom pedagogy. What did protocols of discussion sessions reveal about facets of teacher and/or student participation that could shed light on findings concerning the first and last response paper? The fourth question linked qualitative and quantitative data, asking: Did any of the background factors identified in the pre- and poststudy questionnaire designed for this research project emerge as significant variables and, if so, did these factors correspond to changes or absence of change assessed in students’ response papers?
3. One of the issues explored in a pilot study was whether to show a part of an episode of the television series (i.e., 15 minutes) or an entire episode. Therefore, on one day during the second week, students watched 15 minutes of a 30-minute episode without interruption. In the third and fourth weeks of the course, students viewed the second episode in its entirety. After watching the second episode, students were asked to report on their preferences about seeing a single segment versus seeing an entire episode. There was virtually unanimous consensus that students preferred to watch episodes in full length. They reported that since each episode consisted of at least three storylines involving different people, settings, and events, they would understand the storylines better if they had the opportunity to see more than a couple of minutes of each plot. Students also noted that watching whole episodes better contextualized the different stories and led to a more connected and satisfying viewing experience, especially since the action and the solution of a conflict within each story was often protracted in order to build up and prolong the viewer’s suspense. As a result, students in the main study watched all episodes in full length.
4. Studies by Herron and Hanley (1992) and Herron, Dubreil, Cole, and Cowie (1999, 2000) look at student learning about culture. However, their research does not explore whether students have had the ability to recognize cultural patterns in presented information, modified initial perceptions, or were able to generate new inferences based on it. See also the work of Scott and Huntington (2002, 2007) regarding genre and resulting cultural perceptions.
5. The authors thank Professor Zsuzsanna Abrams, Curriculum Coordinator of the Lower Division Program, for her generous assistance in and suggestions for implementing this project in its formative stages.
6. Students used information available on the Lindenstrasse Internet Web site (www.Lindenstrasse.de). This Web page contains characters’ biographies that identify who the characters are and the relations between them. The biographies also referred to past events that had led the characters to each character’s current positions and situations, and that were referred to in the television episode. Each biography also contained a picture of the character. These biographies were available on the blackboard sites of all fourth-semester German classes.
7. In two of the four classes, these in-class reviews were conducted in English. Both the English and German group discussions followed the same format in both learning conditions. While a small degree of performance difference favored the English-language classes, the general trends reported in this article were consistent for all four classes. For a complete analysis of findings, see Hammer (2008, ch. 6).

REFERENCES

Judith Hammer


APPENDIX A

Directions for Students' Response Papers

- Bitte auf deutsch!
- Bitte auf dem Computer schreiben!
- Bitte mindestens 1,5 Seiten!

1. Briefly summarize what you have seen! No more than 5 sentences!!!

2. Provide at least five more paragraphs about the following:
   a. What are the social and cultural topics and issues that are portrayed? Please describe as many as possible.
   b. What other cultural aspects did you recognize in the episode (e.g., people's behaviors, patterns of everyday life, conflict-solving strategies, etc.)?
   c. What is the personal reaction towards the issues discussed in the episode? Did you expect such topics to be part of German culture? Do you think that this is representative of German culture? Does it help you to understand German culture better? Does it make you like German culture better?
   d. What cultural similarities and differences to your own culture did you recognize?

APPENDIX B

Instructors' Lesson Plans

Lesson Plan, Viewing 1—February 7, 2007—Das Eheversprechen

Before class: Write down the names of the characters on the board. Keep groups of characters together in the following way – 1) Lisa & Murat, 2) Ifi, Jan & Momo, 3) Gabi & Andi, 4) Elena, Vasily & Julian.

Inform students that it is important that they stay on task and work fast during the previewing activities since the episodes are 30 minutes long and since you have to accomplish quite a bit before watching LS.

00-04 Ask students to come to the board and to write down (in bullet-point style) important information about the characters in German. (They can use their homework for this activity.)