Beliefs about assessment and language learning: Findings from Arabic instructors and students

Victoria C. Nier, *Center for Applied Linguistics*
Francesca Di Silvio, *Center for Applied Linguistics*
Margaret E. Malone, *Center for Applied Linguistics*

Abstract

Much attention has been devoted to the positive effects that instructor understanding of the principles and practices of sound assessment can have on learning outcomes and measurement of these outcomes. However, less has been written about the potential benefits of increasing students' understanding of assessment. This paper describes exploratory focus group research conducted with students and instructors of Arabic as part of a project to develop an oral proficiency assessment training resource for this audience. Focus group discussions in response to broad questions eliciting needs and desires for such a

Victoria C. Nier (M.S., Georgetown University) is a Research Assistant in World Languages and International Programs at the Center for Applied Linguistics. She provides professional development in assessment for world language teachers and conducts research with students, teachers, educators, and researchers on meeting their assessment needs. She has experience working with PK-12 and university level educational institutions in domestic, bilingual, and international contexts.

Francesca Di Silvio (M.A., Georgetown University) is a Research Associate in World Languages and International Programs at the Center for Applied Linguistics. She delivers teacher training in world language assessment through both online and face-to-face methods, and has developed assessment and rater training materials in Modern Standard Arabic and Mandarin Chinese. Her research interests include language assessment literacy and language learning outcomes from study abroad.

Margaret E. Malone (PhD., Georgetown University) is Associate Vice President for World Languages and International Programs at the Center for Applied Linguistics. She has more than two decades of experience in language assessment research, language test development, materials development, delivery of professional development and teacher training, data collection and survey research, and program evaluation. Her research interests include language assessment literacy, oral proficiency testing, and program evaluation. Her publications include a chapter in the *Encyclopedia of applied linguistics* and articles in *Foreign Language Annals, Language Testing*, and *Language and Linguistics Compass*, as well as several book chapters and white papers.

resource showed areas of mismatch and areas of overlap between instructors and students regarding beliefs about assessment and language learning. In the context of research on student motivation and the importance of assessment in the learning process, findings suggest that promoting understanding of assessment among language students could support and clarify language learning goals, and thus improve students' overall assessment and learning experiences.

The United States has an urgent need for proficient speakers of languages other than English to meet evolving social, economic, and security demands (United States Department of Education, 2009). Literature in the field of language education highlights the critical role of motivation in language learning (Dörnyei & Schmidt, 2001), particularly the ways in which student expectations for learning and assessment can shape the learning experience and ultimate language learning outcomes (Nikolov, 2001; Schulz, 1996). Research also suggests a vital connection between reliable and valid assessment and effective language teaching and learning (Brown, 2004), and further recommends assessment literacy, or an understanding of the principles of sound assessment, as basic knowledge for all instructors (Boyles, 2005; Inbar-Lourie, 2008; Stiggins, 1995; Taylor, 2009).

This paper describes findings from exploratory focus groups conducted with students and instructors of Arabic as part of a larger project to develop an oral proficiency assessment training resource....

This paper describes findings from exploratory focus groups conducted with students and instructors of Arabic as part of a larger project to develop an oral proficiency assessment training resource for this audience. Broad focus groups questions were designed to elicit participants' needs and hopes for such a resource. The focus group discussions showed a conspicuous contrast in student and instructor beliefs about Arabic assessment and language learning, and these conflicting beliefs were deemed an important consideration in resource development. This paper presents findings from the focus groups in the context of research on student motivation and the importance of assessment in the learning process. Although results from this qualitative study are specific to the small group of informants, they raise important questions about how greater understanding of assessment can influence student and instructor expectations and communication as well as language learning outcomes.

The paper begins with the background for the study including the importance of effective assessment in building a cadre of citizens proficient in world languages. Next, it reviews current research on the alignment of assessment and instruction and student motivation and language acquisition. While it goes beyond the scope of this paper to conduct a review of the rich literature on Arabic instruction in the United States, interested readers are referred to Wahba, Taha, and England (2006) for such a discussion. The paper then describes the methodology of the exploratory research gathered from diverse groups of Arabic language students and instructors and examines results regarding their perceptions of assessment and language learning. The paper concludes with a discussion of possible

Beliefs about assessment and language learning

implications and suggestions for further research on student and instructor beliefs about assessment and language learning.

Background

In recent years, the need for proficient speakers of world languages has been recognized by researchers and policy-makers in the United States including President Obama, who cited the importance of encouraging foreign language skills in a 2011 town hall meeting (The White House, Office of the Press Secretary). O'Connell and Norwood (2007) highlight world language proficiency as increasingly important for national security, international trade, business, government, and legal and medical fields, as well as a requisite for a globally aware population. Jackson and Malone (2009) echo this description of the critical need in the United States for speakers of languages other than English, citing the demands of national security and diplomacy, international commerce and economic development, policies and services for a multilingual domestic population, and global awareness and scholarship at all educational levels. It is clear that the United States requires citizens with high-level language abilities to meet national and international economic, diplomatic, and defense needs. How, then, should the United States address the priority to develop proficient multilinguals?

> ... world language proficiency [is highlighted] as increasingly important for national security, international trade, business, government, and legal and medical fields, as well as a requisite for a globally aware population.

The use of assessment will be a key factor in building this language capacity. Jackson and Malone (2009) discuss the need for a comprehensive national strategy for language learning and emphasize that successful language programs must include systematic, high-quality assessment. Assessment aligned with instruction empowers instructors and students, allowing students to demonstrate their proficiency and instructors to assess and adjust their teaching methods as well as to measure student progress, and thus ensures positive washback (Hughes, 2003; National Education Association, 1983; Shepard, 2000). Assessment is also important because it provides for accountability in evaluating language programs (Norris, 2009). As Jensen (2007) explains, regular and reliable assessment is essential to measuring the effectiveness of language programs and identifying areas for improvement. For assessments to yield maximum positive impact for stakeholders, however, they must be selected and used effectively. Understanding how to select or develop, administer, and interpret assessments requires knowledge of the principles of effective assessment, or assessment literacy (Popham, 2009; Stoynoff & Chapelle, 2005).

While studies suggest the positive effects of developing assessment literacy among language instructors, the benefits of promoting knowledge of assessment principles among language learners have been less researched. The study described in this paper reports results from qualitative, open-ended research on student and instructor beliefs about assessment and language learning. Findings from focus

groups suggest a need to reconcile mismatches in beliefs between students and instructors. It is the hope of the authors that increasing student knowledge of assessment through training resources and other means would have a salutary effect on motivation, goal-setting, and language learning experiences.

Assessment and Language Learning

Many researchers have identified connections between assessment and positive outcomes in the classroom. Instructors who lack formal assessment training may be impeded from implementing effective classroom assessment, however. As such, assessment education is highly recommended for teacher training and professional development. These research findings are described below.

Educators are frequently driven by external pressures to emphasize large-scale summative testing over classroom-based assessment (McMillan, 2003). With this narrowed assessment focus, many in the measurement and instruction communities perceive assessment as fundamentally large-scale, high-stakes, judgmental, and removed from the day-to-day details of teaching (Harlen, 2007). On the contrary, effective assessment is strongly linked to classroom teaching and student learning; in fact, the literature suggests that effective assessment during and at the end of a course may be one of the most important factors leading to student learning and success (Brown, 2004-2005; Bryan & Clegg, 2006; Havnes, 2004; McMunn, McColskey, & Butler, 2004; Popham, 2009; Taylor, 2009; Yorke, 2001).

> *... effective assessment is strongly linked to classroom teaching and student learning; in fact, the literature suggests that effective assessment during and at the end of a course may be one of the most important factors leading to student learning and success.*

To use assessment correctly, instructors must understand the basic principles of assessment design and implementation. Unfortunately, Stiggins (2007) notes that though instructors spend from one-third to one-half of their instructional time on assessment and assessment-related activities, most lack formal training in assessment. A recent study conducted by the American Council on the Teaching of Foreign Languages (Malone, Swender, Gallagher, Montee, & Whitcher, 2009) supports the particular need for language assessment literacy, reporting that many instructors use standardized tests for languages or purposes other than those intended by the test developer. The instructors surveyed also expressed a desire for more information about assessment practices.

When instructors understand fundamental assessment principles, they are better able to use assessment to improve their own teaching and students' learning (Brown, 2004-2005; Popham, 2009; Shepard, 2000). Weigle (2007) and Stiggins (1995) thus recommend that assessment education be a part of the professional development of every instructor. This emphasis on fostering assessment literacy for instructors raises the question: could increased knowledge about assessment be helpful for students? Research on student motivation related to the assessment experience is discussed below.

Beliefs about assessment and language learning

Student Motivation and Language Acquisition

Literature in the field of second language acquisition indicates that motivation is a key factor for success in language learning (Dörnyei & Schmidt, 2001). Gardner (2001) identifies student attitudes towards the learning situation as a critical piece of integrative motivation that affects the ultimate attainment of language learners. Julkunen (2001) describes students' specific motivations to complete a task and further learn a language as variable and highly affected by the learning situation. In examining language learning outcomes, student perceptions of the instructional context, including assessment practices, are clearly an important consideration.

Research suggests that assessment is often a negative and frustrating experience for students, especially when expectations for learning are unclear. Across subject areas, studies report that students find assessment to be a mysterious process, disconnected from course content and uninformed by instructor explanation of intentions (Hodgman, 1997; Kings, 1994). Nicol and Macfarlane-Dick (2006) review studies documenting mismatches between student and instructor perceptions of assessment standards and note that students must understand what a good performance is to be able to incorporate feedback and improve. Examining the effects of different types of feedback on student motivation, Butler (1987) finds that learners attribute success more to effort than to self-worth when receiving task-related comments instead of numerical grades or standard praise; individual comments also led to higher performance and more continued interest. Dweck (1999) notes that students' self-perceptions of learning and ability can lead them to avoid or seek challenges, such that assessment feedback that spurs pride of accomplishment, rather than a focus on fixed standards and comparison to others, can produce greater motivation.

For language learners, frustrating experiences with assessment can have a great impact on motivation and ultimate attainment. As Nikolov (2001) reports, students who perceived themselves to be unsuccessful language learners most frequently cited assessment activities as causes of anxiety and unpleasant classroom experiences. Schulz (1996) suggests that when language learners' instructional expectations are not met, they may begin to doubt the instructor's ability and lose motivation for learning. Still, many studies have documented a mismatch between student and instructor expectations for language learning (Brown, 2009; Chavez, 1997; Kern, 1995; Kuntz, 2000; Polat, 2009).

For language learners, frustrating experiences with assessment can have a great impact on motivation and ultimate attainment.

To make assessment a positive and motivating experience, Black and Wiliam (1998) recommend that students receive clear, timely, and individualized assessment feedback; such constructive feedback encourages active involvement in learning and has a large effect on student self-esteem which ultimately influences attainment. Additionally, as Butler (1987) has shown, task-involving feedback can be more motivating to students than ego-involving feedback such as grades. Sadler (1989) further insists that for students to benefit from assessment feedback, they must assume responsibility for the assessment process through exposure to and

understanding of standards and goals. In the language classroom, Schulz (1996) recommends that instructors and students share course expectations in order to motivate learning. Brown (2009) also supports discussion of learning expectations, and argues that when language instructors share ideas about effective pedagogy, students are more supportive of the classroom activities they are asked to complete.

Given that student motivation and consequently learning can be highly influenced by beliefs about classroom activities like assessment, it is worthwhile to examine whether increasing student understanding of assessment can be beneficial for language learning.

Given that student motivation and consequently learning can be highly influenced by beliefs about classroom activities like assessment, it is worthwhile to examine whether increasing student understanding of assessment can be beneficial for language learning. Research has shown that developing students' understanding of assessment purposes and practices has a positive impact on learning (Rust, Price, & O'Donovan, 2003; Smith, Fisher, McPhail, & Davies, 2009) and that students are open to increased empowerment in assessment (Francis, 2008); these studies do not focus on the language classroom, however. By providing insight into student and instructor beliefs about assessment and language learning, including mismatches between the groups, the qualitative data gathered from Arabic students and instructors in this study suggest an important area for further research.

Methodology

Data Collection

This paper describes the results of exploratory research conducted in the initial phase of development of an oral proficiency assessment training resource for students and instructors of Arabic. Small, open-ended focus groups were planned with target audience members to help determine resource components and structure that would benefit potential users. The research questions that informed data collection procedures were as follows:

RQ1. What is the current level of assessment knowledge of the students and instructors sampled?

RQ2. Do students and instructors express an interest in increased understanding of assessment?

RQ3. Are there mismatches in beliefs, expectations, and goals regarding assessment and language learning between students and instructors?

The researchers conducted four focus group interviews: two with instructors of Arabic and two with students of Arabic from varied educational settings in the United States. Focus groups rely on a group interview technique in which a small number of people (Dörnyei, 2007, suggests six to twelve) participate in a guided discussion on a topic of interest, typically facilitated by a moderator and recorded by a note-taker. Focus groups allow participants to talk to each other as well as the interviewer, and this interaction helps stimulate participant responses,

Beliefs about assessment and language learning

thus generating data that could not be gathered via a written questionnaire or a one-on-one interview (Krueger & Casey, 2000). Focus group methodology was additionally chosen for this study because it allowed researchers to interview multiple participants at once in a limited timeframe for this data collection. Table 1 on the next page shows the composition of the four focus groups conducted for the study. Language background and heritage learner status were self-reported by the focus group participants.

The first instructor focus group included five Arabic language instructors from a large public university (hereafter, State University). One participant was female and four were male, and all were native speakers of Arabic. The second instructor focus group included four Arabic language instructors from three suburban public high schools in a district with a large population of heritage learners (Suburban High School). Two participants were female and two were male, and all were native speakers of Arabic. The third focus group comprised seven students from the Arabic language program at the same State University as the first instructor group. Three participants were female and four were male, and the group included both upper- and lower-level learners and two heritage learners. Although it is likely that these students were taught by some of the participants in the State University instructor group, for purposes of anonymity the exact relationships among participants from these groups were not solicited. The fourth focus group comprised six students from an Arabic language program at an urban public high school (City High School). Five participants were female and one was male, and all were first- or second-year students of Arabic. There were no heritage learners in this group.

Table 1. Composition of Focus Groups

Participants	Program	Language Background	Heritage Learners*
Instructors N=5	Large public university	Native speakers	Some in student population
Instructors N=4	Suburban public high school district	Native speakers	Many in student population
Students N=7	Large public university	Upper- and lower-level learners	N=2
Students, N=6	Urban public high school	Lower-level learners	N=0

*Heritage learner population and individual status as reported by informants

While the researchers had planned to convene a focus group with Suburban High School students in order to gather data from students and instructors in the same program, this was not logistically feasible. There was also no possibility of constructing a City High School instructor group because there is only one Arabic instructor in that setting. Though the particular constellation of focus group settings was not ideal for cross comparisons, the researchers felt it was

broad enough to provide compelling information about student and instructor beliefs. The varied instructional contexts from which focus group participants were drawn incorporate a mix of formative and summative assessments according to individual curricula.

The first three focus groups were led by a facilitator and recorded by a note-taker. The final focus group, conducted with students from City High School, was led by a facilitator who also took notes. Each focus group lasted about one hour, and all were audio-recorded to aid later reconstruction of the data. Participants in the focus groups were compensated for their time.

Each focus group followed a semi-structured format in which participants are given a list of questions to guide discussion but encouraged to elaborate and discuss other topics as they arise (Dörnyei, 2007), with the goal of generating robust data that would not be captured through a strictly scripted procedure. Participants were informed that they would be participating in a focus group about Arabic assessment and language learning, provided with copies of the questions for discussion, and assured that their contributions would remain anonymous. They were told that their responses would guide the discussion, and that it could proceed in a different direction than that of the prepared questions. As such, the focus group moderator allowed conversation to flow unimpeded and did not comment on participant output or survey participants as to their agreement with previous statements.

The focus group questions were designed to elicit broad feedback on student and instructor beliefs about language learning and the assessment process, as well as respond to the research questions. To address Research Questions 1 and 2 regarding assessment knowledge and interest, participants were asked directly about their experiences with language assessment. To address Research Question 3, student and instructor focus group questions were made parallel to allow for direct comparison of responses regarding instructional and assessment experiences and program goals. The following seven questions were used in the instructor focus groups:

1. What are your experiences teaching Arabic? How long have you taught it and in what situations?
2. What are your experiences with testing and assessment in your current Arabic program? What kinds of assessment practices does your program currently use?
3. What are the goals of your current Arabic program? What are your students expected to be able to do by the time they finish the program?
4. What are your experiences assessing the oral proficiency of your students?
5. Have you ever been trained to assess your students' oral proficiency using a large-scale test like the ACTFL OPI, or something like it? If so, what was your experience with that training?
6. What questions do you have about oral proficiency assessment?
7. Is there anything else you would like to tell us about your experiences teaching Arabic, assessing your students' oral proficiency, or a workshop on how to assess oral proficiency, like the ACTFL OPI workshop?

Beliefs about assessment and language learning

Students were asked two additional questions to probe their reasons for studying Arabic and background with self-assessment, as these elements were expected to be addressed in the assessment training resource. The following nine questions were used in the student focus groups:

1. What are your experiences learning Arabic? How long have you studied it and in what situations?
2. What made you decide to study Arabic?
3. What are your experiences with testing and assessment in your current Arabic program? What kinds of assessment practices does your program currently use? Does your Arabic program focus on speaking skills and assessments?
4. What are the goals of your current Arabic program? What are you expected to be able to do by the time you finish the program?
5. Do you feel that you understand the way that your Arabic language skills are assessed?
6. Have you heard of the ACTFL Proficiency Guidelines? If so, what is your experience with them?
7. Have you ever completed a self-assessment of your Arabic skills? What was it like?
8. What questions do you have about assessments of your Arabic language skills?
9. Is there anything else you would like to tell us about your experiences learning Arabic and being assessed?

Data Analysis

An emergent coding process was used to analyze the results of these exploratory focus groups, and researchers used the following data-driven procedures to provide for describable and replicable data analysis. First, typewritten notes taken during each focus group were checked against audio- recordings for accuracy in developing full transcripts. Coding of the transcripts followed procedures outlined in Dörnyei (2007) in line with a system of open coding evolving from the data gathered (Mackey & Gass, 2005). Independent review of the focus group transcripts by two researchers yielded a preliminary list of descriptive codes to represent the relevant themes raised. The researchers compared their emergent coding schemes to create a combined list of codes and then re-coded each transcript according to the consolidated list. Following this second phase of analysis, the researchers worked to resolve any discrepancies in coding. Codes were refined, added, and edited, until a final coding scheme that was determined to most closely reflect the data was developed. The final coding scheme was then applied to the transcripts by the researchers during a third and conclusive pass of coding.

The final coding scheme identified two overarching categories: beliefs about assessment and beliefs about language learning. Within each category there were multiple codes, each representing a belief about assessment or language learning expressed by at least one student or instructor during the focus groups. The following section describes the focus group findings using these coding categories.

Results

Beliefs about Assessment

Table 2 summarizes the beliefs about assessment raised in the focus groups. Marked cells denote that the belief was cited by at least one participant in that group. A plus (+) indicates statements in support of the belief, a slash (/) indicates statements opposing that belief, and a swung dash (~) indicates conflicting statements expressed by different participants within a group. Due to the semi-structured format and small sample size of the focus groups, the number of mentions of a particular belief is not quantified.

Table 2. Beliefs about Assessment

Beliefs	State University Instructors	Suburban High School Instructors	State University Students	City High School Students
Assessment is valid.	+	+	/	+
Assessment is focused on the four skills.	+	~	/	+
Textbooks are disconnected from assessment.	+	+	+	
Instructors need to make assessment materials.	+	+		+
Time and resources for assessment are limited.	+	+		
Assessment training is helpful.	+	+		
Dialect should be assessed.		/	+	

Note: Blank cells indicate that the belief was not discussed in that focus group.

As Table 2 shows, the participating students and instructors did not always concur in their beliefs about assessment. In several cases, students from State University held different beliefs than instructors, including those from the same program. There were areas of overlap between the instructor groups in expressing desire for greater time and resources for assessment and additional assessment training, as well as across student and instructor groups in beliefs that the textbooks used are disconnected from assessments and instructors need to make their own assessment materials. The following paragraphs provide illustrative quotes and comparisons across groups for each belief about assessment raised in the focus groups.

One belief about assessment that emerged across all groups concerned assessment validity. In describing experiences with assessment in their programs, instructors from both State University and Suburban High School expressed confidence that their assessment systems accurately gauge student levels and are

appropriately tied to classroom instruction: "the exams that we give them, the activities that we give them is [*sic*] geared toward strengthening, assessing their level" (State University Instructor Group); "for unit exam tests we still have different tests that we're using based on what we're targeting in class" (Suburban High School Instructor Group). City High School students affirmed that their tests are "designed purposefully," commenting that "the tests help us see what we really know and what we need to work on" and "everything has a reason…it's not there for nothing." Students from State University, however, questioned the degree to which their assessment results reflected real-world proficiency. As one student noted, "I don't really know if what I'm getting in class is really an accurate reflection of what I can and cannot do." It is interesting to note the conflict in beliefs about assessment validity between students and instructors from the same program at State University. While the students' lack of confidence in the validity of their assessments may be due to limited understanding of their use or to fundamental problems with the implementation of a test in a given context, in either case, this questioning of assessment practices could be a demotivating factor for their language learning (readers are referred to Bachman (2005), Chapelle (2012), and McNamara (2006) for further discussion of validity in language assessment).

Regarding beliefs about the focus of assessment in their programs, there were areas of consensus and dispute among the four groups. Instructors from both State University and Suburban High School agreed that assessment should be "testing the four skills of the language" (State University Instructor Group) and "include all the components of the language" (Suburban High School Instructor Group); instructors from Suburban High School dissented within their group, however, on whether that balance was actually reached. City High School students thought that their assessments balanced testing of different skills, though they particularly valued speaking tests: "I like the speaking tests better. I think I learn more when I have to speak it." Suburban High School instructors and State University students expressed a clear desire for more and improved oral assessment: "We ignore to some extent the speaking part, which I want to see more emphasis on" (Suburban High School Instructor Group); "I feel like some of those speaking proficiency… parts of the test aren't really gauging us in terms of our ability to speak" and "I'd just like to emphasize again how helpful it is to be assessed more on speaking than we are currently" (State University Student Group). Again, a mismatch in beliefs about the focus and purpose of assessments between students and instructors from State University could reflect an obstacle to language learning that needs to be overcome.

A prominent line of discussion for instructors from both State University and Suburban High School as well as for students from State University was the limitations of the textbooks used with respect to instructional goals and assessment. State University instructors explained that the book that they use "doesn't help people to be productive" and reported student complaints about its design and content. Suburban High School instructors similarly mentioned student complaints and commented that the available books "don't have anything oral" and "[don't] cover what we want the book to cover." Students from State

University commented on the "gap" between the vocabulary used in the book and that needed for oral proficiency assessment and suggested that the situation be considered "as you guys develop your tests" as well as "in terms of the people that are actually going to be doing the grading of the tests…if it's possible to have a native speaker who understands how college courses are taught in the United States to conduct those assessments." City High School students do not use a textbook in their program.

Another belief about assessment that arose in the focus groups concerned instructors' need to make their own assessment materials. Instructors from both State University and Suburban High School and students from City High School discussed instructor development of assessments and teaching materials. One State University instructor stated that he was never provided with assessments, and in discussing the book, others commented that "as instructors…we also have to provide them with other materials" and "we work outside the book a lot." Suburban High School instructors explained that "lacking the good materials…is one of the serious problems that we are facing" such that "most of the times we make our own or bring in from other resources," especially in the case of oral assessment due to the paucity of books with oral scenarios. City High School Students expressed satisfaction with their program's use of various instructor-developed materials as distinct from the "usual traditional learning of reading from a book."

Two other beliefs about assessment were frequently cited by instructors from both State University and Suburban High School. First, multiple focus group participants brought up the challenge of limited time and resources in developing and administering assessments, noting that "one of the most difficult things is weighing the time and dividing the time between teaching skills and at the same time get enough time to test" (State University Instructor Group) and "we need the money to do [our own] resources, we need rich resources" (Suburban High School Instructor Group). Second, discussion within both groups indicated a clear regard and desire for assessment training among participants: "training really situates the instructor in a position where he or she knows what to expect from the students" (State University Instructor Group); "I'd like to see workshops that Arabic instructors are invited to where they would go over stuff like this we can utilize in the classroom" (Suburban High School Instructor Group).

A final belief about assessment that generated considerable discussion among students from State University was the desire for testing of dialect, which points to a much-discussed issue in Arabic instruction, the details of which are beyond the scope of this paper (see Wahba, Taha, and England (2006) for further, recent discussion of this issue in the field). Participants in this group talked about the lack of testing in dialect and how the availability of such tests would encourage a more serious study of dialect, commenting that "it would be great to be able to know where you stand on dialect because that's what you really use when you're talking to people" and "it would make sense to base [assessment] on whatever would be most commonly understood by a person on the street." The issue of rating dialect use in assessment was raised by an instructor from Suburban High School, a district with many heritage learners, who questioned how she should

Beliefs about assessment and language learning

grade students who use dialect and mentioned that dialect was not accepted in assessments used to test out of the language requirement.

Beliefs about Language Learning

Table 3 summarizes the beliefs about language learning raised in the focus groups.

Table 3. Beliefs about Language Learning

Beliefs	State University Instructors	Suburban High School Instructors	State University Students	City High School Students
Proficiency goals are clear.	+	+	/	/
Instruction is focused on the four skills.	~	/	/	+
Students are motivated.	~	~	+	+
Students need to learn with a textbook.	+	+		/
Students believe Arabic is difficult.	+	+	+	+
Students learn more studying abroad.	+	+	+	
Students wish to compare instruction.			+	+

Note: Blank cells indicate that the belief was not discussed in that focus group.

As Table 3 shows, there were incongruities in beliefs about language learning between instructors and students regarding proficiency goals, focus of instruction, student motivations, and the need for a textbook. There were some areas of overlap among the groups, however, particularly in assertions that Arabic is considered a difficult language to learn and that students learn more Arabic when studying in an Arabic-speaking country than in the United States. The following discussion provides illustrative quotes and comparisons across groups for each belief about language learning raised in the focus groups.

A belief about language learning discussed in all four groups concerned the clarity of program proficiency goals. Instructors from both State University and Suburban High School articulated clear proficiency-related goals for their Arabic programs: "I would just summarize the goal of our program here is to be able to teach our students to be independent in the language" (State University Instructor Group); "they're really using survival Arabic skills at that level" (Suburban High School Instructor Group). A State University Instructor explained, "at the beginning of each semester, we put our objectives before the students…it's a contract between us and the students." Participants in the student focus groups were less certain about proficiency goals, however. One student from

State University commented that as a senior he was frustrated with his level of language learning given "inflated" expectations from freshman year, and another noted that "other than lessons as markers of progress, I feel like we're not given anything else in the classroom in terms of this is what you should be able to achieve by the end of the semester." While students from City High School readily communicated their understanding of their program's language learning goals, they did not explain these goals in terms of proficiency. As one City High School student stated, "I think the goal is just to have fun learning Arabic…not as much as having standards to meet." In another case of student-instructor mismatch, it is noteworthy that State University students questioned the clarity of their program's proficiency goals while State University instructors felt that these goals had been firmly stated.

There were areas of disagreement within and across the four groups regarding beliefs about the focus of instruction, as with beliefs about the focus of assessment. One State University instructor commented that "there's more focus on… speak[ing] but there isn't a lot of emphasis on the other skills as well," while others in that group asserted that "I personally can't imagine the oral proficiency just standing by itself…rather than tying in to the other three skills" and "nobody can argue that not teaching the alphabet is something that's fine." A Suburban High School instructor, by contrast, noted that while language instruction must involve all four skills, "in our daily teaching I would say we focus on reading and writing." State University students expressed a desire for more speaking instruction and practice: "I definitely echo the sentiment that…speaking should be a more regular activity in classroom settings." Finally, students from City High School believed that, as with assessment, their program's instruction balanced the four skills: the instructor "makes sure that we learn the language from all different angles" and "it's not just if you can read it, it's not just if you can write it, it's not you can just listen to it—no, it's everything that goes into learning a language and knowing the language." The different beliefs expressed by students and instructors from State University on the focus of instruction with regard to oral skills highlights a potential frustration to the language learning process for all involved.

Another belief about language learning discussed in all four groups touched on the diversity of motivations animating the study of Arabic. Students from State University mentioned employment incentives; heritage, political, travel, and religious interests; and general interest in the language as reasons for studying Arabic. City High School students enumerated career and academic reasons and desire to learn the language as motivations for language learning, and were particularly attracted to the unfamiliarity of Arabic. Every City High School focus group participant noted how Arabic stood out from the typical languages offered in high school, describing it as "interesting," "exotic," "weird," and "completely different." Instructors from both State University and Suburban High School listed a variety of student motivations for language learning including those cited in the student focus groups, but also discussed lack of motivation among some students: "I'm telling you the real thing they speak about when you ask them…why you choose Arabic to study, because this is what they offer" and "some students are not

motivated" (State University Instructor Group); "sometimes somebody is there by chance" and "when you ask them why are you here…they would say, because my parents want me, not because he likes the language or because he wants to communicate or it's his native language" (Suburban High School Instructor Group). While self-selection for the voluntary focus groups could be a factor in explaining discrepancies between student and instructor accounts of language learning motivations, it is nevertheless important to note these differences between the groups.

The belief that American students need a textbook to learn language was strongly expressed by instructors from both State University and Suburban High School despite their reservations about the textbooks used in their programs. An instructor from State University was explicit that "no student in the United States is going to study anything without a textbook" and another explained, "a lot of students, if they don't have a textbook, they feel they don't have a reference, and if they don't have a reference, they feel like things are sort of haphazard and ad hoc." As a Suburban High School instructor noted, "I need a book, parents want a book." In direct contradiction to these assertions, City High School students, who do not use a textbook, were pleased with this "unstructure" as "it's not helpful to learn something out of the book because then you can't really use it in the outside world to really communicate with people." Students in this group did not question their lack of a textbook but wondered how other programs effectively teach with a textbook: "Would a person be more interested in the language if they learned it from a book or without a book?" and "Did anyone ask how they teach it with a book?" Though the need to work with a textbook was not directly raised by State University students, participants in this group did express frustration with their book's disconnect from real-world language use. State University instructors' insistence that students feel a textbook is necessary for language learning may therefore be another example of a mismatch between beliefs of students and instructors in the same program worthy of further explanation.

One belief about language learning shared by participants in all groups is that American students view Arabic as a challenging language. Students from both State University and City High School stated that Arabic is more challenging than other commonly taught languages: "I don't know why they don't have Arabic and Chinese and other really difficult languages when they start you off in high school" (State University Student Group); "Spanish is something very easy to learn and I wanted something that was a challenge" (City High School Student Group). Instructors from both State University and Suburban High

One belief about language learning shared by participants in all groups is that American students view Arabic as a challenging language.

School concede that Arabic can be hard to learn, especially the script: "they're comparing it with Spanish, and the reason they think it's difficult is because they would rather learn Arabic in Latin code" and "it is difficult, I know, if someone is writing from left to right all his life" (State University Instructor Group); "it's not an Indo-European language like the rest of the languages, so it is difficult I have

to admit" and "the writing part is the challenge in Arabic because of the changing shapes of letters" (Suburban High School Instructor Group). The instructors go further, however, to say that misconceptions about Arabic can hinder student language learning if left unchecked, commenting that "the first thing that we do as instructors is debunk these myths; if somebody says, well, Arabic is very hard, so you say…Arabic is very systematic" (State University Instructor Group) and "marketing of Arabic language is needed" to encourage potential learners who hear from other students and counselors that Arabic is difficult (Suburban High School Instructor Group).

A final belief about language learning shared by instructors from State University and Suburban High School as well as State University students is that students learn best in study abroad environments. Multiple State University students expressed this belief, commenting that "studying in Egypt was a lot more both intense and helpful than studying in the United States" and "I learned more in that short time immersed in the program [in Yemen] than I did here." A State University instructor explained that "we tend to encourage the student who has some kind of talent in the language to immerse in bigger communities…or to go overseas for some time," and the Suburban High School instructor focus group closed with the comment, "I would love to see some exchange programs—the best way to learn the language is to go to the country where they speak it." Within the student focus groups, final questions raised by the participants focused on how their learning compared with other programs: "I was kind of worried-slash-curious to know the level of rigor of the Arabic programs at other universities" (State University Student Group); "I would like to know, are they quizzed the same ways we are, do they do speaking tests, or do they have multiple choice?" (City High School Student Group).

Discussion

In response to Research Question 1, discussions during the focus groups illustrated that the current level of assessment knowledge of the students and instructors sampled is unequal and could be improved. Though the scope of the study is limited and findings may not be generalizable to other populations, the focus group discussions suggest that promoting assessment knowledge among students as well as instructors could benefit both groups and potentially lead to improved learning outcomes.

The focus group results indicate that instructors in both focus groups had higher levels of assessment knowledge than the students who participated in data collection. This finding is not unexpected, as instructors are more likely to have experience and training in assessment practices and purposes. The data show that instructors believe the assessment techniques they use are valid and that assessment training is helpful. These beliefs reflect participating instructors' awareness of the need to develop their own assessments and the limitations of time and resources they face in implementing effective assessment.

As demonstrated in general discussion during the focus groups not specifically addressed in this paper, students in both groups had less fundamental knowledge

of assessment than the instructors sampled, which could contribute to the lower confidence they express about assessment. The State University students were unsure that their assessments were valid, and repeatedly commented that they did not trust the real-world utility of the assessment approaches they had experienced. Though the City High School students had faith in their instructor's assessment practices, they were curious to learn more about other types of assessment. Students in both groups also expressed opinions about how their assessment systems could be improved, through more assessment of speaking (State University students and City High School students) and of dialect (State University students). This questioning of assessment purposes by students and recognition by instructors of the challenges of scarce time and resources to develop and implement assessments seems to recommend increased assessment knowledge for both groups.

In response to Research Question 2, both students and instructors showed an interest in increased understanding of assessment. Some student participants questioned the purpose and validity of the assessments they experienced, suggesting receptivity to and desire for greater understanding of assessment, while instructors in both groups reported a desire for more assessment training.

Participants in the instructor groups stated that assessment training was important in providing instructors with the information necessary to create reasonable expectations for student learning outcomes. Further, the instructors requested more and continued assessment training that can translate to effective classroom practice, including training specifically targeted to the challenges commonly faced by instructors of Arabic, such as guided practice, sample tasks, and other resources to support instructors in creating their own assessment materials (given the lack of materials available in the language).

Student participants also indicated an interest in increasing their knowledge about assessment without using the term assessment training. Students in both groups displayed uncertainty regarding the proficiency goals of their Arabic programs and raised questions about the tests being used to assess their progress towards those goals, which could reflect an issue with the tests or with student understanding of their objectives. They expressed a desire to understand what their assessment results meant in terms of functional language ability outside of the classroom, as well as in comparison with other language learners. State University students in particular were keenly aware of the difficulties facing Arabic learners who must use a variety of dialects and registers to communicate effectively. Assessment training would help students find and, more importantly, trust in the answers to these questions by inviting them to engage with the assessment practices used in their programs and in the greater Arabic language learning world.

It is worth particular mention that in both student focus groups, the concerns raised about assessment related to test validity in assessing "real world" language abilities rather than to grades. During study design, the researchers sought to ensure that focus group protocols did not make value judgments or mention evaluations or grades. Is was therefore interesting to discover that students at both the high school and university levels were curious about the functional utility of assessment in providing information about their proficiency instead of its effect on

their grade point averages. This result from a self-selected group of learners echoes the conclusions of Butler (1987) and Dweck (1999) regarding the motivating influence of feedback focused on potential improvement rather than rank, and may suggest that knowledge about language proficiency and how it is assessed is specifically desired by language learners.

Finally, in response to Research Question 3, there were many instances of mismatches of beliefs, expectations, and goals regarding assessment and language learning between students and instructors, including representatives of the same university program. In terms of beliefs about assessment, students were unsure about the validity of their assessments, while the instructors felt no such doubts. Students also questioned the balance of assessment focus among the four skills that was promoted by the instructors, and unequivocally expressed a desire for testing in dialect while instructors were unsure how to address this issue. In terms of beliefs about language learning, while instructors felt that the proficiency goals of their programs were clear, students were unsure about those goals. Students also questioned the balance of the four skills in instruction that was claimed by the instructors. Although the instructors were suspicious about some student motivations, all of the student participants described themselves as very motivated learners. Finally, instructors were certain that all students needed and demanded a textbook in order to learn Arabic, while students demonstrated that this was not the case.

These areas of mismatch in instructor and student beliefs about assessment and language learning could pose a challenge to learner motivation if students feel their expectations are not being met (Schultz, 1996), which in turn could threaten learning outcomes. It is hoped that resources for instructors and students designed to build understanding of assessment and encourage dialogue about assessment theory and practice could help to address such areas of mismatch and thus bolster student motivation.

Conclusion

Based on this small qualitative study, it seems advisable to promote understanding of assessment among language students and instructors to address the interest in assessment knowledge demonstrated by the study participants. The robust focus group discussions were critical in informing development of oral proficiency assessment resources for instructors and students of Arabic. As a result of the understanding of the beliefs about assessment and language learning of both groups and limitations in their knowledge gained from this study, the resources were designed to include explanations of the importance of oral proficiency, discussion of the use of dialects in instruction and assessment, testimonials from students about their how their oral proficiency ratings translated to real-world use, and numerous samples of student target language responses to prompts at various proficiency levels, among other features.

The exploratory character of this research limits the conclusions that may be drawn from it, however. Given the short time frame and limited resources for this phase of study, only four focus groups were conducted and only two of the

groups sampled instructors and students from the same program. Larger numbers as well as additional methods of data collection would have provided more robust results and avoided possible artifacts arising in the data due to idiosyncrasies of the programs sampled, such as the comments about textbook use. In future research investigating areas of mismatch between students and instructors, a wider sampling of both groups from various institutions would be desirable, as well as collection of course instructional and assessment materials to allow for triangulation of data.

Ideally, the enhanced assessment understanding fostered by the newly developed oral proficiency assessment training resources would provide both students and instructors with a fundamental understanding of the purposes of different types of assessment including specific assessments used in their classrooms. Improved assessment knowledge could increase the confidence of both audiences in assessment as part of the learning process by positioning assessment as a way to mark progress towards real-world goals. Further, it might provide a bridge for dialogue between students and instructors about reasonable expectations for language learning and alleviate areas of misunderstanding between the two groups. Such outcomes have the potential to lead to increased student motivation and greater language learning.

It is the authors' view that further research on student response to assessment training resources could prove illuminating. It is clear that assessment plays a fundamental role in language learning, and that student motivation, student and instructor expectations, and communication between both groups are intertwined and critical in shaping learning outcomes. Investigating how an increased understanding of assessment and communication of expectations between students and instructors influences learner motivation and language attainment, as well as the best ways to encourage such effects, would be a fruitful area for future research.

References

Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly, 2*(1), 1-34.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Boyles, P. (2005). Assessment literacy. In Rosenbusch, M. (Ed.), *National assessment summit papers* (pp. 11-15). Ames, Iowa: Iowa State University.

Brown, A. V. (2009). Students' and teachers' perceptions of effective foreign language teaching: A comparison of ideals. *Modern Language Journal, 93*(1), 46-60.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* White Plains, NY: Pearson Longman.

Brown, S. (2004-2005). Assessment for learning. *Learning and Teaching in Higher Education, 1*, 81-89.

Bryan, C., & Clegg, K. (Eds.). (2006). *Innovative assessment in higher education*. New York: Routledge.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest and performance. *Journal of Educational Psychology, 79*(4): 474–82.

Chapelle, C.A. (2012). Validity argument for language assessment: The framework is simple… *Language Testing, 29*(1), 19-27.

Chavez, M. (1997). Students' and teachers' assessments of the need for accuracy in the oral production of German as a foreign language. *Modern Language Journal, 91*(4), 537-563.

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

Dörnyei, Z., & Schmidt, R. (Eds.). (2001). *Motivation and second language acquisition.* Honolulu, HI: University of Hawaii Press.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* Philadelphia, PA: The Psychology Press.

Francis, R. A. (2008). An investigation into the receptivity of undergraduate students to assessment empowerment. *Assessment & Evaluation in Higher Education, 33*(5), 547-557.

Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1-19). Honolulu, HI: University of Hawaii Press.

Harlen, W. (2007). Formative classroom assessment in science and education. In J. McMillan (Ed.), *Formative classroom assessment* (pp.116-135). New York: Colombia University Teachers College Press.

Havnes, A. (2004). Examination and learning: An activity-theoretical analysis of the relationship between assessment and educational practice. *Assessment and Evaluation in Higher Education, 29*(2), 159-176.

Hodgman J. (1997). The development of self- and peer-assessment strategies for a design and project-based curriculum. *UltiBASE.* Retrieved from http://ultibase.rmit.edu.au/Articles/dec97/hodgm1.htm

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*, 385-402.

Jackson, F., & Malone, M. (2009). *Building the foreign language capacity we need: Toward a comprehensive strategy for a national language framework.* Washington, DC: Center for Applied Linguistics.

Jensen, J. (2007). National foreign language policy: A state language coordinator's perspective. *Modern Language Journal, 91*(2), 261-264.

Julkunen, K. (2001). Situation- and task-specific motivation in foreign language learning. In Z. Dörnyei & R. Schmidt (Eds.) *Motivation and second language acquisition* (pp. 29-41). Honolulu, HI: University of Hawaii Press.

Kern, R. G. (1995). Students' and teachers' beliefs about language learning. *Foreign Language Annals, 28*(1), 71-92.

Beliefs about assessment and language learning

Kings, C. B. (1994, November). *The impact of assessment on learning.* Paper presented at AARE Conference, Newcastle, Australia. Retrieved from http://www.aare.edu.au/94pap/kingc94179.txt

Krueger, R. A., & Casey, M. A. (2000). *Focus groups: A practical guide for applied research* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Kuntz, P. (2000). Beliefs about language learning: Students and their teachers at Arabic programs abroad. *African Issues, 28*(1-2), 69-76.

Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Malone, M., Swender, E., Gallagher, C., Montee, M. & Whitcher, M. (2009). *Survey of assessment uses and needs.* Unpublished final report.

McMillan, J. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice, 22*(4), 34-43.

McMunn, N., McColskey, W., & Butler, S. (2004). Building teacher capacity in classroom assessment to improve student learning. *International Journal of Educational Policy, Research and Practice, 4*(4), 25-48.

McNamara, T. (2006). Validity in language testing: The challenge of Sam Messick's legacy. *Language Assessment Quarterly, 3*(1), 31-51.

National Education Association. (1983). *Teachers' views about student assessment.* Washington, DC: Author.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nikolov, M. (2001). A study of unsuccessful language learners. In Z. Dörnyei & R. Schmidt (Eds.) *Motivation and second language acquisition* (pp. 149-169). Honolulu, HI: University of Hawaii Press.

Norris, J. M. (2009). Understanding and improving language education through program evaluation: Introduction to the special issue. *Language Teaching Research, 13*(1), 7-13.

O'Connell, M. E., & Norwood, J. (Eds.). (2007). *International education and foreign languages: Keys to securing America's future.* Washington, DC: The National Academies Press.

Polat, N. (2009). Matches in beliefs between teachers and students, and success in L2 attainment: The Georgian example. *Foreign Language Annals, 42*(2), 229-249.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education, 28*(2), 147-164.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Schulz, R. A. (1996). Focus on form in the foreign language classroom: Students' and teachers' views on error correction and the role of grammar. *Foreign Language Annals, 29*(3), 343-364.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29.7,* 4-17.

Smith, C., Fisher, R., McPhail, R., & Davies, L. (2009). *Student assessment literacy in the Bachelor of Business.* Retrieved from http://www.griffith.edu.au/__data/assets/pdf_file/0008/225773/Good-Practice-in-Assessment_Case_Study_BBus.pdf

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-245.

Stiggins, R. J. (2007). Conquering the formative assessment frontier. In J. McMillan (Ed.), *Formative classroom assessment* (pp. 8-28). New York: Colombia University Teachers College Press.

Stoynoff, S., & Chapelle, C. A. (2005). *ESOL tests and testing: A resource for teachers and program administrators.* Alexandria, VA: TESOL Publications.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29,* 21-36.

United States Department of Education. (2009). *Consultation with federal agencies on areas of national need.* Retrieved from http://www.ed.gov/about/offices/list/ope/iegps/consultation-2010.pdf

Wahba, K. M., Taha, Z. A., & England, L. (2006). *Handbook for Arabic language teaching professionals in the 21st century.* Mahwah, NJ: Lawrence Erlbaum Associates.

Weigle, S. C. (2007). Teaching writing teachers about assessment. *Journal of Second Language Writing, 16*(3), 194-209.

The White House, Office of the Press Secretary. (2011). Remarks by the President in a town hall meeting in Atkinson, Illinois [Press release]. Retrieved from http://www.whitehouse.gov/the-press-office/2011/08/17/remarks-president-town-hall-meeting-atkinson-illinois

Yorke, M. (2001). Formative assessment and its relevance to retention. *Higher Education Research and Development, 20*(2), 115-126.