



Final Report

Expert Panel on Community Interpreter Testing and Certification

sponsored by the
Interpreting Stakeholder Group
of the
Upper Midwest Translators and Interpreters Association

Plymouth, Minnesota

June 13-15, 2007



prepared by

Bruce T. Downing, Ph.D.
and Scott Homler, Ph.D.
Program in Translation and Interpreting
College of Continuing Education
University of Minnesota – Twin Cities

Table of Contents

EXPERT PANEL OVERVIEW		4
BACKGROUND AND CONTEXT		10
Summary of opening remarks by Bruce Downing		10
Panel participants and state representatives		12
PLENARY PRESENTATIONS		13
Lessons learned from national interpreter certification initiatives and recommendations for spoken language community interpreter certification based on these experiences		
Laurie Swabey	Registry of Interpreters for the Deaf	13
Roseann González	Federal Court Interpreter Certification [not available]	
William Hewitt	National Center for State Courts – Consortium for State Court Interpreter Certification	21
Lessons learned from interpreter certification initiatives at the state level and recommendations based on these experiences		
Elizabeth Nguyen	California Healthcare Interpreting Association	27
Maria-Paz Avery	Massachusetts Medical Interpreters Association	33
Hungling Fu	Washington State	42
Lessons learned from corporate and non-profit certification exams		
Janet Erickson-Johnson	Language Line Services [not available]	
Frances Butler	NetworkOmni Multilingual Communications	48
Roseann González	National Center for Interpretation, U of Arizona [not available]	
Jean Turner	The California Endowment/Hablamos Juntos	55
Lessons learned from current state efforts (summary report)		
Enrica Ardemagni; Indiana; Maria Michalczyk; Oregon; Armando Villareal, Iowa; Carol Berg, Minnesota		61
QUESTIONS POSED TO WORKGROUPS		65

WORKGROUP REPORTS	66
I. The Relationship between Individual States and National Initiatives	66
<i>Group members: Armando Villareal, Maria Michalczek, Mauro Yanez, Carol Berg, and Enrica Ardemagni (Berg and Ardemagni, facilitators)</i>	
II. Interpreter Competencies and Test Development & Administration	67
<i>Group members, Interpreter Competencies: Karin Ruschke, Isabel Arocha, Elizabeth Nguyen, Roseann González, Laurie Swabey (Bruce Downing, facilitator)</i>	
<i>Group members, Test Development and Administration: Hungling Fu, Jean Turner, Frances Butler, Nataly Kelly (facilitator)</i>	
III. Implementation	71
<i>Group members: Maria-Paz Avery, William Hewitt, Shiva Bidar-Sielaff, Janet Erickson-Johnson, Cindy Roat (Patricia Ohmans, facilitator)</i>	
<i>(For concluding discussion, see “Expert Panel Overview,” p. 9)</i>	
<hr/>	
APPENDIX A: AGENDA	78
APPENDIX B: RECOMMENDED READINGS	79
APPENDIX C: PARTICIPANT BIOGRAPHIES	82

Expert Panel on Community Interpreter Testing and Certification

-- Overview --

[This summary of the Expert Panel was published separately online in July 2007 and remains available at <http://www.cce.umn.edu/creditcourses/pti/downloads/index.html> as "June 2007 Expert Panel Brief Report." A short oral report on the Expert Panel was presented in a plenary session at the October 2007 IMIA Conference, preceding the National Certification Panel; copies are available from the author.]

Bruce T. Downing

A meeting important to the field of interpreting took place in Minnesota on June 13-15, 2007. The "Expert Panel on Community Interpreter Testing and Certification" was organized in order to further an ongoing national discussion of certification for health care interpreters. This intensive and very productive three-day meeting of 13 invited language professionals was held at the Radisson Conference Center in Plymouth, Minnesota.

The Expert Panel was sponsored by the Interpreting Stakeholder Group, a membership group within the Upper Midwest Translators and Interpreters Association, which held its annual conference immediately following. It was organized by the University of Minnesota's Program in Translation and Interpreting in cooperation with Century College and funded by a generous grant from the Bush Foundation of Saint Paul.

The Expert Panel had three main goals:

- to convene a group of people with experience and expertise regarding assessment of interpreter qualifications;
- to begin to assess what we know and what we need to do to build a fair and reliable certification process;
- to explore how state and national initiatives can work together for their mutual benefit.

Those invited to participate on the panel included a number of experts in language testing and psychometrics, most of whom had been involved in the creation of existing tests or screenings for interpreters. Others represented membership organizations in the field of interpreting, such as the California Healthcare Interpreters Association (Elizabeth Nguyen), the International (formerly Massachusetts) Medical Interpreters Association (Izabel Arocha), the National Council on Interpreting in Health Care (Shiva Bidar-Sielaff), and RID Inc. (Dr. Laurie Swabey). In addition to the panel itself, individuals were invited to report on state initiatives in Indiana, Iowa, Oklahoma, and Oregon, four states which have taken significant steps toward certification or regulation of interpreters; Carol Berg represented Minnesota's Interpreting Stakeholder Group.

The first day and a half of the Expert Panel was devoted to presentations on a number of existing programs for interpreter testing and certification, emphasizing lessons learned. Dr. Laurie Swabey summarized the history of RID certification of ASL interpreters. Dr. Roseann D. González outlined the process she used in the development of the original Federal Court Interpreter Certification exam. William Hewitt

of the National Center for State Courts reviewed the history of the Consortium which has developed and coordinated the administration of state court exams. Both Dr. Maria-Paz Avery and Elizabeth Nguyen reported on stages in the evolution of the exam originally developed for the Massachusetts Medical Interpreters Association and piloted in California with the cooperation of the California Healthcare Interpreters Association. Dr. Hungling Fu reported on the medical interpreter certification administered by the State of Washington.

The invited state representatives (Dr. Enrica Ardemagni, Maria Michalczek, Armando Villareal, and Mauro Yanez) gave updates concerning developments in their states since the publication of Cindy Roat's *Certification of Health Care Interpreters in the United States: A Primer, a Status Report and Considerations for National Certification* (The California Endowment, September 2006).

The second morning, Dr. Frances Butler talked about the proprietary exam used by NetworkOmni Multilingual Communications, and Janet Erickson-Johnson described the development and the nature of the Language Line Services exam, available nationally through "Language Line University". Dr. González reported on the Medical Interpreter Competency Examination (MICE) designed at the University of Arizona. Finally, Dr. Jean Turner of the Monterey Institute of International Studies described a medical interpreting examination which has been developed in California with the involvement of Hablamos Juntos and with support from The California Endowment. (Panelists not mentioned elsewhere included Cynthia Roat and Nataly Kelly, both independent consultants, and Karin Rushke, President of International Language Services Inc.)

During the final day and a half, the panel first discussed broad questions concerning the certification process. Then the panel was divided into three workgroups to address the following topics:

- a) state to state and state-national coordination of efforts;
- b) interpreter competences, test design, and levels of certification;
- c) implementation: how to take the next steps in moving toward certification.

Finally, each workgroup reported its ideas back for discussion by the whole panel, leading to a series of recommendations, which can be only partially summarized here.

1st Group's Recommendations: Coordination of State and National Efforts

Note: "State" does not necessarily mean a public or government agency; it may be a state interpreter organization, advocacy group, or advisory committee.

How can states work with the national initiative?

1. States can use common definitions for words like "registry."
2. States can share a model for "registries" (software) developed by CHIA.
3. States can compile list/descriptions of currently available training, and make recommendations regarding what should comprise training.
4. States can serve as pilot sites for training, certification tests, etc.
5. National body could centralize/coordinate state initiatives.
6. National body can compile a national registry of certified interpreters in each state.

7. National body could provide model statutes or legislation.
8. National body can recommend a model training program, drawing ideas from existing state programs.
9. National body can help states avoid duplicating efforts, e.g. to recruit trainers, to have a rotating train-the-trainer institute.
10. National body can offer language to help states build a “business case” for interpreters to show how trained interpreters are cost-effective.

How can states work together?

1. States may organize individual state databases or registries of interpreters. A “registry” may vary in complexity from state to state. In some cases it may just be a database of unverified information. A standard definition of the word “registry” is a listing of people who meet some minimal qualifications who are thereby authorized by the state to interpret or to have their services reimbursed.
2. States can develop training programs that can inform national training standards and efforts in other states.
3. States can get started in preparing interpreters to eventually become certified.
4. States should continue to meet face to face (as in this forum).
5. States could join together to sponsor train-the-trainer institutes.
6. States can develop tests in various areas or for different languages.
7. States can collaborate on research to build a business case – or compile such research that exists.

2nd Group’s Recommendations: Interpreter Competencies and Test Development

The workgroup on Interpreter Competencies and Test Development agreed that standards, training and certification must be thought of together. They also agreed that what is needed is not just an exam but a certification *process* and that it would be important to have a “battery” of tests—not one test but a series. The group reached a preliminary consensus on potential components for the certification process.

- (Recommended) A reliable assessment of general language proficiency in two languages. The exact sequencing and form of proficiency testing was left for further discussion.
- A minimum amount of training in an approved program prior to testing.
- A test (given in either language) of professional standards and ethics in which scenarios would be presented verbally. This might be administered during or following training that would use a standard training module.
- Skills Test I: A role-play scenario to test consecutive interpreting skills using primary care situations, probably delivered via a recording.
- Skills Test II: Sentence conversion in both directions, to test specialized vocabulary and register (in provider speech and patient speech) and sentence conversion skills in a wide range of clinical settings.
- Skills Test III: Sight translation (using relatively brief and non-legal texts).
- Skills Test IV: Simultaneous interpreting role play **OPTIONAL**—for those who seek an “endorsement” for simultaneous skills.

Important Next Steps

The Competencies and Test Development workgroup identified the following action items that would be important to undertake in order to prepare for development of a national certification process:

- Conduct a detailed Job Analysis of health care interpreting. First, a decision will need to be made as to how to define the field: does it include home health visits, dentistry, physical therapy, chaplaincy, mental health assessments, etc.
- Conduct a review of the National Standards of Practice and other standards publications, to determine which items are reflected in the Job Analysis, and which items from the Job Analysis reflect a lack of concordance with the standards, so that these items might be addressed through training and testing components (e.g., sight translation of legal documents, as described above)
- Conduct surveys to determine most common content domains for various language groups, and those that are common to all interpreters, in order to select content domains for role plays.
- Conduct a legislative review to ensure that the certification process reflects applicable law.
- Hold focus groups/expert panels to determine the specific competencies, i.e., the knowledge, skills, abilities and tasks (KSAT), to be performed by professional medical interpreters, and therefore evaluated. Conduct a review of literature and gather empirical data, such as job descriptions, information on error analyses, etc.
- Draft a list of desired test preparation materials (or develop materials that would be needed).
- Draft a checklist of “steps toward certification” to help prospective candidates prepare themselves.

3rd Group’s Recommendations: Implementation of a Process: Next Steps

The Implementation workgroup proposed a series of steps for moving toward national certification, with emphasis on the first six months (through December 2007).

Step One in the process is to identify a national organizing or coordinating group to take the lead. A next step is to make available a full report from this Expert Panel to inform future efforts. Then it is essential to secure funding in order to proceed further. With initial funding in place, or promised, the organizing/coordinating group will need to identify a panel of experts in medical interpreting that would then begin to do or to oversee the sorts of tasks that the Competencies and Testing workgroup identified.

For the period after six months the Implementation workgroup offered a list of other tasks that would need to be accomplished, such as compiling existing information on certification issues, what job analyses already exist, and what we know about adequate proficiency levels. There will need to be a process to select the approved certifying body. But even before there is an actual test, there will need to be a template for test design, and there will need to be opportunities to get feedback on that, and then on a draft test for a particular language pair, and so on—a continuous iterative process. So a process needs to be put in place for the steps of design, testing, and review. Another task will be exploring ways to market the certification process to stakeholders: getting interpreters and the medical establishment to buy into the value of certification, and to understand the lengthy iterative process. Along the way it will be necessary to secure funds for the actual test development, and for tests in multiple languages. To make the process collaborative and to obtain buy-in, there’s a need for additional steps, such as conducting forums, obtaining and analyzing survey feedback, again furthering the iterative process. The workgroup stopped at this point (since this meeting was after

all just a first effort in designing and elaborating a possible process), but obviously there's much more to be done.

In the meantime, while this process is hopefully moving forward, other related activities will also be underway. The development of a set of national standards for health care interpreter training and education is being planned by the National Council on Interpreting in Health Care (NCIHC), and it seems likely that at the state and local level both governmental and non-governmental organizations will be developing some of the products envisioned: rosters and registries, guidelines for recruiting, language screening programs, interpreter training/education programs, and perhaps a template for language access legislation—a model statute that could be proposed in state legislatures.

The Implementation workgroup suggested a need for national oversight of the process: identifying a group that can immediately begin the tasks of organizing, convening, getting money, coordinating efforts, disseminating information, and assuring that the process moves forward in a more coherent way than any of the scattered past efforts. But in this model the oversight or coordinating organization would NOT be making judgments about competencies to be tested, or test design, or the eventual administration of certification testing as they are organizing the process. Their responsibility would be coordinating what needs to be done for test development; they would be bringing representatives of stakeholder organizations on board for collaboration and communication; they would be convening the committee of experts on medical interpreting and competency testing. It is THAT group--the committee of experts--that would need to make some decisions. And the representatives of state organizations and other stakeholders—people who are supporting the effort—will also need to be participating in the decision-making. There will need to be subcommittees, such as the group of representatives of organizations (and other stakeholders), and the group specifically responsible for test development. There will be a need for regular communication among all the parties involved, which might be the responsibility of a designated subcommittee --- creating and managing a listserve, for example.

Considering various possible options for getting the process going, the implementation workgroup recommended that the National Council on Interpreting in Health Care (NCIHC) be called upon to do the initial organizing and fulfill the coordinating function—to convene an ongoing expert panel and a task force with stakeholder representation that would make the decisions. The recommendation of the NCIHC was based on the fact that the NCIHC has a recognized national leadership status and that it has been conducting forums on certification at major conferences across the country for several months. It has immediate possibilities for grant funding for this particular initiative, having been approached by a major foundation to prepare a grant proposal. If funding is obtained, its project would be for the NCIHC to convene a group of stakeholders who would decide in a stakeholder meeting how the effort should be organized and who should be on the committee of experts, what subcommittees or related workgroups would be needed, and what tasks would need to be assigned to different groups.

Closing Discussion.

Three theoretical options had been laid out by the Implementation workgroup, as follows:

- Option 1: Form a new coalition with representation from CHIA, IMIA, NCIHC etc.
- Option 2: Ask states to form their own stewardship coalition.
- Option 3: The NCIHC coordinates and secures funds for the collaboration.

There were no advocates for Option 2; discussion centered around options 1 and 3. An argument in favor of Option 1 was that it seemed more collaborative. The idea that the NCIHC was the only natural choice to take the lead under option 3 was also questioned.

The discussion recognized that competition for leadership was a potential obstacle to progress. Communication and cooperation are needed so as not to end up with parallel efforts and competing certifications. If any one organization is going to take the lead, no matter what that organization is, there obviously has to be a lot of outreach to all other stakeholders. But, it was suggested, perhaps the present duplication of effort, and potential competition, is happening because there presently IS no coordinating agency. And the Implementation Workgroup wanted everyone to understand clearly that the NCIHC is not proposed as the decision-making body, only as an existing organization that is prepared to undertake the initial step of sending out the invitations and convening the coalition. They argued that an independent coalition (Option 1) would be a new organization, not immediately positioned to receive funding and start things moving.

Yet it was countered that forming a coalition of equal partners would not necessarily entail a new organization. It could be a group of organizations that come together and sign agreements to work together on a common task, communicating with each other and perhaps dividing up the work. One possibility, following this idea of a coalition with no designated “leader,” would be that the NCIHC could get the funding and other members of the coalition play other roles, such as publicity. There was no question though that however the coalition is organized, some organization needs to secure and share funding, because without funding nothing can be done.

The Minnesota Expert Panel ended with this discussion, having made important progress on the major issues it was asked to address but without full agreement on how to take the next steps.

###

Background and Context

Summary of Bruce Downing's opening remarks on Day One of the Expert Panel

The Expert Panel on Community Interpreter Testing and Certification was planned by the Interpreting Stakeholder Group and funded by the Bush Foundation of Saint Paul, Minnesota, as part of the Linking Voices project designed to improve and expand interpreting services throughout Minnesota. The project was administered through the College of Continuing Education, University of Minnesota, with the support of Century College.

Goals for the Expert Panel

- To convene a group of people with experience and expertise regarding assessment of interpreter qualifications
- To assess what we know and what we need to do to build a fair and reliable certification process
- To explore how state and national initiatives can work together for their mutual benefit.

Some considerations

- We can learn a lot from what has been done already. Still--
- We are at the beginning of a process; we'll mainly be talking about next steps.
- We are not a formal organization and we are not here to form an organization.
- We can make recommendations and hope others will read our report.
- In some cases we may only be able to identify and clarify the questions.

Let's begin with the broad questions

- Should there be a national certification?
- Should there be a general certification or just medical (& workers comp & child protection & special ed, & mental health & ...?
- Should the passing level be minimal competence or mastery?
- How should education, training, and testing be fitted together?
- Does certification testing imply licensure?
- What do we know? Where do we start?

Format for this Expert Panel

- Reports on what has been done and lessons learned
- Plenary discussion of lessons learned and how to identify next steps
- Workgroup sessions on four main areas
- Reporting back for general discussion and recommendations

Workgroups

- A. Interpreter competencies
(essential knowledge and skills to be assessed; level(s) of performance)
- B. Test development and administration
(assessment methods, design, validation, levels, language-specific factors)
- C. Implementation
(how to move forward, how to get input and buy-in, how to structure the process)
- D. State-level processes and cooperation

Some essential definitions

“**Credentialing** is the umbrella term that includes the concepts of accreditation, licensure, registration, and professional certification.”

“**Professional certification** is the voluntary process by which a non-governmental entity grants a time-limited recognition and use of a credential to an individual after verifying that he or she has met predetermined and standardized criteria.”

“**Licensure** is the mandatory process by which a governmental agency grants time-limited permission to an individual to engage in a given occupation after verifying that he/she has met predetermined and standardized criteria . . .”

“**Registration** [in one of 3 basic meanings] is the governmental process by which a governmental agency grants a time-limited status on a registry, determined by specified knowledge-based requirements (e.g., experience, education, examinations), thereby authorizing those individuals to practice, similar to licensure.”

--all definitions from *The NOCA Guide to Understanding Credentialing Concepts*, NOCA, 2005.

Panel Participants

Izabel Arocha	President, International Medical Interpreters Assn.
Maria-Paz Avery	Educational Development Center, Newton, MA
Shiva Bidar-Sielaff	Manager of Interpreting Services, UW Madison
Frances A. Butler	Independent language testing consultant
Janet Erickson-Johnson	Language Line Services
Hungling Fu	DHS, State of Washington
Roseann D. González	Professor and Director, NCITRP, U of Arizona
William Hewitt	retired, National Center for State Courts
Nataly Kelly	Independent consultant
Elizabeth Nguyen	CHIA Board Member
Cynthia Roat	Independent consultant
Karin Ruschke	President, International Language Services, Chicago
Laurie Swabey	RID; Director, CATIE, College of St. Catherine
Jean Turner	Professor, Monterey Inst. of International Studies

Host: Carol Berg, Interpreting Stakeholder Group & UCare Minnesota
Organizers: Bruce Downing, University of Minnesota
Veronica Newington, University of Minnesota
Tara Gibbs, Century College
Facilitator: Patricia Ohmans, Health Advocates

State representatives:

Enrica Ardemagni - Indiana
Carol Berg - Minnesota
Maria Michalczyk - Oregon
Armando Villareal - Iowa
Mauro Yanez - Oklahoma

Plenary Presentations

Lessons learned from national interpreter certification initiatives and recommendations for spoken language community interpreter certification based on these experiences.

RID Certification: Lessons Learned

Laurie Swabey, Registry of Interpreters for the Deaf

Good morning. I am Laurie Swabey, and I will be talking about RID certification. I am the chair of the Interpreting Department at the College of St. Catherine, where we offer a B.A. degree in ASL/English interpreting. Our college is one of five regional centers in the nation to be awarded federal funding to work on interpreter education, and one particular initiative our Center is leading is medical interpreting. So I have a particular interest in this topic, although I am not a testing or certification expert. The RID office would have loved to have sent someone; however they are planning our national convention in August, expecting over three thousand members, so the national office is very involved with the national conference right now, but wishes you very well.

So, I am giving more of a member perspective. I have been a member of the RID since 1976 and have been certified by the RID since 1977 and teaching interpreting since 1980. My interest and expertise is really in interpreting education, not testing and certification. But I think I have several things I hope will be of use to you in your work, and I hope that stakeholders in the field of sign language interpreting and the field of spoken language interpreting can work collaboratively to advance both fields. In developing certification processes for sign language interpreters over the years, we have certainly made our share of mistakes, and I hope you can benefit from what we've been through. As a side note, my involvement with spoken language interpreting started in 1990, when I met Bruce Downing and I began doing some teaching at the University of Minnesota in their program.

I thought I'd start off with a little history. RID was established in 1964 and as Bruce was saying, we actually started out as a registry of interpreters. Interpreters were considered "on the registry" if they were recognized by two other members as being competent and ethical. It wasn't until 1972 that we actually started a testing process. And at the time we thought that was a great thing. As we look back, we see we went from a system where people over time had to prove their competence both with professionals and with community members, to a system where it was just one snapshot of what people could do at one particular moment. So even though we were sure of ourselves at the time, we can now see how our first system was not as bad as we originally thought it was.

In 1972, when we started this formal testing procedure, we had about four hundred members and now we have over 12,000 members. Those are not *certified* members, but those are *members*. There has been some concern in our organization now that we have more non-certified than certified members. In 1972 when testing started, we had no degree-granting programs for interpreting, which those of us in

education saw as an issue. In 2007, we have approximately ninety-nine associate degree programs, and although this number changes, I think we may be up to about thirty-three B.A programs. We also have one master's degree in sign language interpreting at Gallaudet University and a master's degree (which is very exciting) in teaching interpreting. This is an online program offered at Northeastern University, and it's in its second year. The first cohort will be graduating at the end of this year. This is going to have an incredible impact on our field, to have a master's program that prepares teachers to teach sign language interpreters, so that's a very exciting piece.

Talking about history, one of the mistakes that we made had to do with our partnerships with stakeholders. We originally started out being in very close partnership with NAD, which is the National Association of the Deaf. In 1972 when we lost some of our national funding that brought us together with them, we went our own separate way, developed our testing system, and did not include the NAD. In the late 80s and early 90s, NAD quite frankly was fed up with the RID and thought that we weren't addressing the issues of stakeholders. And they decided to establish their own certification process. I don't think that was good for the field, but I think it was a big wake-up call to the RID that we needed to be more inclusive and to consider stakeholder input. And after a lot of negotiation and meetings and planning, the NAD and the RID did join forces and created a new test, which we started administering about two years ago. It was named the NAD-RID National Interpreting Certification. I did bring information about the certification process and test, if any of you later during a break would like to see it. This is our practice DVD. It's very important to have test materials for people to practice, and this is very similar to the actual test, so I did bring that if anybody would like to look at that. I also brought examples of what our application materials look like. You may also be interested in this, our Journal of Interpretation. Again, another important component of having a testing or certification process is the component of education.

Another piece that we learned (and it took us a while to learn this lesson) was to have realistic expectations of what certification can accomplish. It's not a cure-all. We originally thought, and I'm speaking very generally, that testing would be a gate that would keep unethical and unqualified people out of the profession. Of course we see people who can pass the test. That they did well at that date doesn't mean that they're ethical. It doesn't even necessarily mean that they're competent in all the situations we would like them to be. So that's been a hard lesson for us to learn, that testing does not fix everything, and it's not everything that we want it to be. It measures people in a particular situation on that particular day. I don't know if you have this in your communities, but there was a lot of political pressure. I think for a long time, we really tried to please everybody; we didn't know how to address those issues, and it's taken us a long time to get a good sense about that.

A couple of other things I'll talk a little bit more about later are again being realistic about testing, and that going along with testing, we found we really needed a grievance procedure and a continuing education component. One of the things we feel is most important is to be able to be a good ethical decision maker and that's not something we can test. We can test people to see if they seem to understand the party line and if they can give a good description of what they would do in a situation and why, but are they truly ethical on the job? We really can't test for that. We've gone back and forth for years discussing *what is the standard?* Are we looking for the gold standard or are we looking for the basic driver's license that is going to get people on the road?

What we have finally come to is a basic professional standard. What we have now for our generalist test is a minimum competency that people need to function in the field.

We have gone other ways before getting to this point. When we had a national standard that people were having difficulty meeting, many states developed their own state tests. At the time I was living in the state of New Hampshire, and we were not having interpreters pass the national test, even though they were graduating from interpreting programs. They were not passing the national test as quickly as we would have liked. The state developed a test, and the idea which a lot of us were in favor of was that people would take the state test as a first step and then go on to the national test. What we found – a broad generalization in our field – is that people stopped. They got their state certification and did not go on to meet the national standard. So we are at the point now of thinking that this basic professional standard on the national level is the way that we should go.

Education is a very important component, both before and after testing. As you have seen from these statistics, we are going towards bachelor's degree programs in interpreting. We have a new provision that is going to go into effect in 2008, that interpreters will not be able to take the interpreting exam, NAD-RID National Interpreter Certification (NIC), if they do not have an associate's degree. So that's beginning in 2008. Right now, anybody that walks in the door can take the test. Starting in 2008, they will have to have an associate's degree. It does not have to be in interpreting; it just has to be an associate's degree. In 2012, candidates for certification will have to have a bachelor's degree, so that's something new that we have gone to.

The other thing that we have found, of course, is that we have to be very clear on the web site and through other avenues to show what our matrix is, what categories we are looking for, how people are being evaluated, how they are going to be assessed, and how they can be prepared. Providing preparation for certification is a big business. People often want that before they go in, even though they may have graduated from a two-year or four-year interpreting program. When and where they received their education, the type of work and continued education they have undertaken since graduating and their comfort with test taking, are all factors that influence their perception of readiness for certification.

Let me just talk about one more thing with respect to interpreting and education. We are very fortunate in interpreter education and with sign language interpreters to have an organization called CIT. The Conference of Interpreter Trainers is an association that brings together more than two hundred interpreter educators every two years for a national conference as well as publishing a newsletter and a proceedings. What is most important now, our biggest step forward, is that, based on the standards developed by CIT, there is now an independent accreditation process to accredit interpreter education programs. And programs are starting to go through that. Currently there are about ten programs that are in the two-year process which includes a self-study and a site visit. So we see accreditation as a very important piece.

Now, I'll talk about the knowledge test. The first thing that our certification candidates have to do is take a knowledge test. It consists of 150 multiple choice items, and there are two dates nationally when it's given. It's given the first Saturday in June and the first Saturday in December. Other than that, people can arrange to take it individually at other times, but it is more expensive.

I want to go back for a moment. It's a little confusing. We have a generalist test, which is really where the focus is, and in the generalist test, you take a knowledge test and then a performance test. We also have a legal certification test. For the legal certification, you have to earn the generalist certification first. And then you take a written legal test, and then a performance test for legal certification. We also have something called oral transmission, which is only of interest to a very small number. It doesn't involve sign language, but only the use of the face and mouth to do oral translation for people who don't use sign language. That was a very heavy political piece in some ways too. We also have an educational interpreting certification that was not developed by the RID or NAD. It was developed by an outside organization, but it is now under the umbrella of RID. One of the concerns with that educational certification is that people called themselves certified, but they were only tested in interpreting in a K-12 setting. However, they may be working in a variety of community settings. When RID first started, we were going to have lots of specialty tests. In fact, we had a performing arts certificate. Now we look back on it, and we say with medical, legal, and educational settings needing qualified interpreters, why did we go from generalist to performing arts? We don't offer that certification any more, but it was something that we did in the past. We also don't offer a medical interpreting certification.

The other piece I wanted to address is that we do have a certificate called the CDI (Certified Deaf Interpreter) which is for deaf interpreters, interpreters who *themselves* are deaf. They often work as intermediary interpreters, as language specialists. That is a separate test, and one that I want to talk about related to the knowledge test, because some of the deaf people who want to take this test are native ASL signers, others are native English users. Some may have become deaf when they were five. Others, born deaf, with deaf parents, have ASL as their native language. Some deaf people are having a difficult time passing the written knowledge test and there was a demand to offer it in ASL. That now is being offered. So, the knowledge test for the deaf candidates is offered both in written English and in ASL on DVD. Candidates view the test questions on DVD and then mark their answers. Interestingly enough, having that piece has not increased the passing rate. And again it is an interesting discussion because if you're an interpreter, you need to be fluent in both languages. So that has been sort of a debate or a topic, whether or not that was a good way to go. I'm sure that's something some of you might face in terms of English-language proficiency for some of the less common languages.

We also have a performance test, of course, and for the performance test candidates for certification watch a DVD. On this DVD, test candidates are asked questions about how they would handle certain situations. They respond to these questions and their answers are recorded on video. Then they also do actual interpreting, simultaneous interpreting. On the old tests, the test before this one, the pass rate was about sixty per cent. Now, the pass rate is about thirty-three percent. On the written test, the pass rate is about eighty-eight per cent. Many candidates pass the knowledge test before or right after they graduate from their interpreting programs. They seem well prepared for this portion of the test, as you can see from the pass rate. With the legal written test, I think about eighty-eight per cent are passing the written test, and also a high number are passing the performance test.

A piece we have found very important is our certification maintenance program (CMP). Interpreters maintain their certification through a CEU program. It's not a lot of

CEUs, in fact it only comes out to about twenty hours a year, but we have found that it really encourages the number of professional development opportunities. You have people that are just chair-warmers at professional development opportunities, but you also find many interpreters are truly engaged, and this has tended to move our field forward.

I will just say that we used to have a very formal kind of grievance procedure with a board making a decision. We have gone to a mediation system. So we are really trying to resolve grievances through mediation, which we have found to be fairly successful.

I will go on to costs. Our current test cost one million dollars to develop, so it was very expensive. The sense that I'm getting from interviewing people in RID is that as much as we would like to have a certification, say in medical or something like that, it probably will not be performance-based, just because of the cost. But it could be that interpreters would achieve the generalist certificate and then take a written knowledge test, perhaps to demonstrate knowledge of the healthcare setting. Anyway, currently test candidates pay high fees to take the test. It costs between five hundred and six hundred dollars to the participant. And that just barely pays for the test. That does not incorporate any profit. We have to have raters who are very skilled at rating the test. When we have a rater training, it costs between fifty to seventy thousand dollars to train fifty raters. It is estimated that maybe thirty of them will rate for us.

We have a burn-out rate at testing centers. Our college is a testing center. People do it from their heart. They are paid twenty-five dollars to administer a three-hour test. They are going in, changing the tape during this three hour period, so they *can* do other work but they are interrupted throughout the testing time. So the administration of the test is heavy in terms of cost and people power. The RID would like to farm out the test. They can't find anybody that's willing to take it on because of the costs and the administrative needs.

Another issue is that it's difficult in some ways to keep the test current. We are still giving a test that was designed in the 1980s. It will be phased out within the next two years. There are things on that test that aren't current any more. When people are talking about technology, it's just simply not current. Test takers are having a hard time even understanding it if they are talking about certain things. So there is the cost of keeping the test current which has been incorporated into this. This is our first test that the RID feels is legally defensible, and reliable and valid. Again, such a test is much more expensive to develop. But of course, that's the direction we needed to go.

Also, it's very expensive when you need to bring in all of the experts (and it sounds like some of you have the expertise), such as psychometricians and the other true experts in testing and certification that need to be involved, as well as stakeholders, which are also a very important part of the process. The portfolio is a way that we have thought about going, although we haven't done it yet, where people make examples of their work on video, have letters of recommendations, and have their degrees. There's been talk of this both in the United States and Canada with their interpreting associations. That's not a way that we've gone yet, but it's certainly something that's under discussion. And as I told you, there is a BA requirement that will be in effect in the near future. Some people believe that there should also be an option for people to waive this requirement by submitting a portfolio. No decisions have been made on this

but interpreters who have been interpreting for thirty years and don't have a bachelor's degree are advocating for a portfolio option or some other option in lieu of the bachelor's degree.

I want to close by saying again that this information on certification is from the perspective of somebody who is a longtime member and an educator. And I think that there are several critical components to keep in mind when considering a certification process: The educational component, knowing what standard you're going for; the grievance procedure; the continuing education procedure; and having realistic expectations about what certification will do for the field.

Q: I'd like to know what your knowledge test exactly tests?

A: There are questions about linguistics on a very basic level, knowledge about culture on a very basic level, knowledge about professional issues and the professional organization, and then interpreting issues. And some of them are just recall questions, others require a higher level of analysis or processing. We have a test prep book that gives examples of all of them.

Q: And then is that used as a screening test to be able to take the performance test?

A: Right. You can't take the performance, unless you've passed that test successfully.

Q: I'm curious. Is it considered in the ASL interpreting community that being able to read and write English is an important skill for ASL interpreters?

A: Yes. And we have talked about having an English proficiency test. There is no plan for that now, although we hope that the A.A. degree requirement, and then the B.A. degree requirement will cover at least part of the English proficiency piece. But no, we have never tested for that. In fact, to get into a lot of interpreting programs historically, you have not had to have proven competency in either ASL or English beyond whatever the college requires for English. So it's an area of weakness for us.

Q: I'm also curious where you got your million dollars, so I can go talk to them about giving us a million dollars.

A: I don't know where we got the million dollars. I really don't. Although maybe it's partly from our membership dues.

Q: Actually, [with regard to] the screening test issue, did you have a lot of discussion (in a field where certified people are hard to come by) about the worries about screening people out that could pass the performance test?

A: No. We have a very high pass rate because most interpreters do go through an interpreter education program. That seems to be a piece that we're very successful at doing. There's a suggested reading list and practice examples.

Q: One of the big concerns in court interpreting in particular is that the pass rates are so low. I mean, we're talking four percent or twenty percent, you know, sort of very low. And

so, there you have a very high level of concern about telling somebody that they couldn't do the performance test because they did badly on a multiple choice written test.

A: That's probably more similar to our deaf interpreters who have a very low passing rate on the written test. And there is concern about that.

Q: I was fascinated. Actually, I'm trying to figure out the issue of why native ASL speakers would have trouble reading English. And then I realized well of course, it's a completely different language. And who's going to teach it? Can you talk a little bit about, just for general interest, how you teach reading written English to an ASL native speaker or language person?

A: Through bilingual, bicultural education programs. And deaf education is a whole other huge issue. Some deaf people have had opportunities to become fluent in ASL and English, either through their school or home environment. Other deaf people may not have had the opportunity to become fluent in ASL and/or English because of the education that was available to them.

Q: It just never occurred to me that native deaf people might not be able to read the newspaper. It just never occurred to me.

A: One of the big issues for deaf people of course is if they've grown up in a home with deaf parents who are fluent in ASL, they've got a first language base. It's much easier for them to learn a second language. If they grew up in a situation where they entered the educational system without any language base, that's much more difficult.

Q: Can you talk a little bit about the practical side of delivering the test? I always hear that there is a lot of wait time, that people are not able to take the test, so that they can, you know, get there. And I know you've been working on improving that. How is that?

A: I will talk about that from the point of being a center, and it's a little bit of a sore point for us because RID used to say that our center was a "test on demand center," which we agreed to, and which means that people can call up and schedule a test. It's not just on a fixed schedule like the first Saturday of every other month. Instead there are a variety of time blocks on various days that are available on a first come, first served basis. What happened was people thought they could call up and say, "I want to take my test on Wednesday at 8:00 a.m.," and that's not the case at all. So that's where some of those complaints come from, that people think that they can call up and schedule a test at a moment's notice. Also, there are limits to the number of test slots, depending on the number of tests administrators. The test administrators who give the test are practically volunteering. The other thing that happens is that people want to wait until the last minute. If their employer has said, "You have to be certified within two years," people wait until they've been employed one year and ten months, and then they try to schedule a test. And when we say we can't schedule you within the next two months because it's summer and our testing center is already full for those limited slots, then they're also unhappy. We're seeing some of the testing centers across the nation close because it is very time intensive. With this new test, there are several different parts, and the person administering the test may have to change or start the tape every ten minutes over a two to three hour period. So it's very disruptive for the person who's also trying to work. From our perspective it's very clearly laid out where you register, when you register, and where you pay. Candidates for the test don't pay us, they pay the national office. But

people don't always read everything and so they sometimes get frustrated. I think it is a good process but that people think that they can just do it instantly.

Q: One other question I had was about the state certification or registries. Have you been successful in moving away from that or do they continue to want to use those state certification processes and registries they had?

A: That varies. The state of Texas has had one for many years. Kansas has one. Wisconsin has one. They vary greatly, and that's a very politically hot issue that I'm not sure I want to touch on in the thirty seconds I probably have left. Some of them are well done and contribute to the field. Others, not so much.

Q: My question was just related to the comment you made about the fact that the national test had a very low pass rate, and states started creating their own certifications at a lower, more basic level. I'm just wondering, this new NIC, how does it fair?

A: This new NIC is what has been determined as the basic standard for entering into the profession.

Q: But how does it compare to those states? I mean, did it end up becoming more basic to address that need?

A: OK. Let me say one other thing about that which I didn't say. With this test, you can pass at three different levels. It's a generalist test with the following levels: the NIC level, the NIC Advanced level and the NIC Master level, so it does take in those three levels. The other thing about this is that it's a valid and reliable test, where the previous test had some issues, and so it was very difficult to compare it to state certification. I can't really address how this compares to state testing. That's your question, isn't it? I think what the national decided was that at least if people could get certified with this, it would meet a basic professional standard that was based on a reliable and valid test.

Q: On the registry, you say there are some certified members and non-certified members. What is required for someone to get onto the register without certification?

A: We no longer have a registry, even though we're still called the RID. Anyone can join as an associate member, and then if you're certified, you can join as a certified member. If you take the test and you're not a member of the RID, you pay a higher fee for the test than if you're a member.

Federal Court Interpreter Certification Examination: Premier Model for Interpreter Testing [presentation not available]

Roseann González, Professor and Director, National Center for Interpretation Testing, Research and Policy, University of Arizona

The Consortium for State Court Interpreter Certification

William Hewitt, National Center for State Courts

My job is to talk to you this morning about the Consortium for State Court Interpreter Certification. And that will be the last time that I give you the full name, the unlovely name, and I'll just be referring to the Consortium itself. So what is it? It was founded in 1995 by the National Center for State Courts, and that followed research I had done in the early 1990s looking for models of interpreter programs around the country to try to help the state courts. And it was in the early 90s that interpreting emerged as a huge administrative issue, and the short version of what I found in my research was there really were no models. There were no model programs to speak of. There were some models of good testing programs, and the first one right off the bat of course was the federal examination that has been described to you in great detail and which I've had the privilege of administering up until my retirement for about five, six years. And there were at that time two states in the country that were doing a good job with testing. Those states had the luck to get a mandate from their legislatures and some money behind it to develop certification programs. There were two states, and that was Washington and New Jersey. There were two states, Minnesota and Oregon who had recently gotten a mandate to do testing and they asked, "What? How are we supposed to do this?" So, they came to me and at that time, the conditions were right. I saw an opportunity based on the research that I had done.

Several of the conclusions that I had reached involved establishing some sort of standard national testing program because test development was so expensive. If basically interpreting in Utah is the same as interpreting anywhere else, why keep reinventing the wheel? Why keep spending these enormous amounts of money? So I used Minnesota and Oregon's money to convene a series of meetings with representatives from Minnesota, Oregon, New Jersey and Washington to sort of hammer out [details]. It immediately emerged that we were not just going to be sharing New Jersey or Washington tests with Minnesota and Oregon. Immediately we understood that we were really setting a national standard test. So, once those early negotiations were over, we established the consortium. Basically the consortium is just a voluntary collaboration between state courts to share money through agreements. There's an actual ten or twelve page agreements document that I've developed, that people sign, and it has all of the rules for being a member of the consortium. It is available online.

The original focus of the consortium was directly on certification testing. I'm using the word "certification" loosely here. Basically the focus was on the development of a standard national exam. The idea was to develop one exam. The states themselves use a standard exam, and the certification process is a *state* process. I make that distinction. If a necessary criterion for certification is to pass this test, it is not sufficient. In most states there might be other things that the state piles on, requirements that have to do with things like criminal history record check. It might have to do with some educational requirements, paying money, and so on. Once the consortium was established, it evolved really into a best practices network.

What are the conditions? Initially, it was a one shot deal, based on the percentage in the state of home speakers of languages other than English according to

the Census Bureau figures. Most of the states were required to put in twenty-five thousand dollars. For a few smaller states, it was fifteen, and for some larger states, California, Texas, a few others, it was fifty thousand dollars. [Consortium states] were required to appoint an official representative and to abide by (and this is very important) the administration standards. You can have a nice instrument and screw it all up with administration. The states needed to participate in governance activities.

So, what are the advantages of having a national standard test? Obviously, you can get published test documentation, credibility, and maintenance of statistics. A standard test allows you to have interstate reciprocity. If somebody passed the test in Utah, you don't have to spend the money again to test them in New Jersey. We've discovered that it contributes to administration innovations, even with the federal exam. By the way, the model for the state test is the federal exam. Basically, the federal government had relatively enormous amounts of money to invest in this pioneering. The states had very little money. Following the process of not reinventing the wheel, the model of the federal exam has been applied in the states.

Q: Why didn't you just use the federal exam?

A: Well, because it's proprietary. I did do a formal inquiry at some point about that. You would look at three different scoring levels to pass. You could in theory do that with the federal exam, but there was just no way that they were going to risk that enormous investment in the test. Remember the federal courts are just a minuscule part of the action in the legal environment in the United States. The exposure of the test in the state structure would just be enormously greater, [with] all of the risks of compromising test integrity. In fact, here's a funny story. This happened in one of our states, and it's relevant to one of the implementation issues I was going to talk about. That has to do with staff turnover. A person was appointed to be the program manager for the interpreter program in a given state. He knew it all, he didn't need to come to training at the National Center or pay attention to what we had learned over the years, and what the standards were. So, the first time he administered the test, by paying the fee, by right, he got to have copies of the exam and administer it. When the exam scoring is over, the scoring is done on a test script, a written document. The test raters mark on that when they hear the scoring unit, and they have their little marks and so on. When the scoring was over, so few people passed, which was typical, this guy mailed out, with the test results form, the summary results form! He actually mailed out the marked copies of the test script! And he just gave the test back to them. So it completely compromised a version of the test that cost thousands of dollars. And it cost that state. I had to actually charge them. I think they got away with eighteen thousand dollars or something like that, which they paid.

The consortium test is in these fourteen languages, and the little numbers in parentheses indicate how many versions of the test we have. We have four different forms in Spanish, and two in Vietnamese. I put a question mark by Laotian because one of the problems when you move out into lesser-spoken languages is where to find the experts. Who said they were experts? What do they know anyway? And so on and so forth. Well, our Laotian experts have died. You know various things have happened, and I don't even know what we would do if we had to administer a Laotian test. Who would rate it?

Start-up issues are really important. One of the main issues that I had to face, and I was just dealing with four people, was the dysfunctional sort of concerns of individuals. A zero sum mentality went into the negotiations about what this test should be like. [For] a couple of personalities it was, "I've got to come out a winner in this. I've got to get something out of this." That was really dysfunctional and it increased the cost of doing this enormously. There were people who had strong ego concerns: "I know the right way to do this. If you do it that way, you're not going to be doing it right, and I need to control this process."

There were genuine problems [and] security concerns obviously. How are we going to manage all that? And in the case of legal interpreting, there were variations in the testing context. In New Jersey, the test was given as a screening exam and for people who were going to be employees of the court and [who] would have administrative supervision and mentoring opportunities. They would have opportunities to learn legal language. I mean, learning the technical vocabulary is child's play here. You know, we all do it every time we go into a new context. You learn the terms really quickly. That's easy to teach. Teaching interpreting, that's a whole other thing. Well, anyway, in New Jersey, the people were going to go into a structured employment setting. In Washington State, the test was basically saying to a judge anywhere in the state, this person I'm sending you is basically competent to interpret right now. So there was a big tension over how much technical language there would be in the exam. One of the advances made with the federal test, changes or evolutions over time, is a more highly specified set of instruction standards in terms of the scoring units: how many grammar units, how many register units, how many modifiers, what technical language. And those are very highly structured [standards]. So at least, from one form to another, remember the forms have to be interchangeable, there's very simple, easy face validity right off the bat.

What are some of the evolving concerns about the consortium? By the way, lest I forget, it will probably come out, but the challenges that I faced at that time were far fewer than what you face, because nobody was doing [this work]. In the medical interpreting arena, you've got all sorts of initiatives going on everywhere, and all these competing proprietary sorts of issues are really going to get in your way. Who's going to win? And of course there's a lot of money technically at stake here. Immediately one of the initial concerns we ran into right away was the importance of having a registry of who took the test and how they scored. Much more quickly than we expected, we got "state hoppers." For example, people took the test in Utah and failed it. So then they ran over to another state and took it again, hoping to be exposed to the same form. We had to restrict their exposure to the form, so they couldn't just keep hurling themselves at the gate and eventually passing it.

Actually, the success of the consortium itself and its wide expansion has proven to be a problem. When we first got going, everybody who started in on it understood the rules, understood the values, understood the administrative challenges, and then as we spread out, centrally we couldn't keep track of it all. We had high staff turnover, and we had things like this incident in this one state. I mean, how could you imagine that in a training session, you would want to go up to somebody and say, "Oh, by the way, don't mail the test out to forty people around the United States." I mean how would you even consider training on that point? So staff turnover was a big problem and central administration was a problem. And that's all of my prepared remarks. So, I have time for questions.

Q: I want to go back to why the states want a state court certification instead of just requiring the federal certification? Is there a difference in the job done by federal court interpreters and the state court interpreters?

A: This is a legitimate issue. The federal test pass rate has been uniform basically at around four to five percent. Now that's of all the people that present themselves for the written test. There's about a twenty percent pass rate on the written test. Of those people who qualify for the oral test, the overall is about four percent. For the federal test, there were no prerequisites whatsoever. "I went to Mexico and spent four weeks one time, spent a lot of time in the bar. I think I can be an interpreter." So, why not use the federal test, or why not use federal certification? Well, you have to get 80 per cent of the scoring units correct. And this is a process intending to be objective. It approximates objectivity. So remember now this is twenty percent of these interpreting challenges you can get wrong and still pass the test, even in the federal [test]. In the state context, we did the same thing. "Well, what should a passing rate be?" And it came out to seventy percent, so you can misinterpret close to one third of the units and still pass the test. But it's the scoring standard. There are enough interpreters for the federal courts now, but the gap in terms of having enough *certified* interpreters is just huge. And that's the issue. You have to squarely confront that issue. What is our standard of testing? What are we trying to accomplish?

Q: You had said earlier that you modeled the state tests on the federal tests. Do you use that same system, doing the language proficiency test first and then the performance test?

A: No. We did not. Initially, we went right at performance. The bottom line is, you know, I don't care how well you do on a multiple choice written exam that allows you to choose the correct choice for prestidigitation. You know, I don't really care about that. What I care about is, "Can you get up and interpret in an interpreting context and conserve meaning?" And there are some other complex things about that. Part of it being that the cost of developing a written test is huge in terms of legal defensibility and so on. And so the issue is what's the bang for the buck? It's true that we replicated in a paper and published a paper on the predictive validity of the written exam. It's absolutely true that there is a significant statistical correlation between passing the written exam and passing the oral exam. As a practical matter, how good a job is it doing? Reasonable people can disagree. And in one of the studies that we did, we did come up with false negatives. In other words, we did the same thing, allowing people to take the oral test who had failed the written test. And we found just enough people who failed the written test and passed the oral test to make us nervous. In an environment (this is for the federal test now) where you're really hungry for people who have face competence, that false negatives thing makes one a little nervous.

Q: With regard to the process, as you mention, the states use the same tests, but the processes are state driven. You mentioned some of the different requirements: education requirements, background checks. Has the national center compiled any information about what the different states are doing and how it's affecting their ability to have a group of qualified interpreters?

A: Well, the first part is yes. If you go to the National Center's website and work your way into the interpreting section under the consortium, you'll find a lot of resources there now, one of which is surveys of what the various requirements by the states are for certification. So yes, you can go and get information about what it takes in Utah versus what it takes in California. But in terms of saying, "Well is this one better than that one?" My personal opinion, based on a lot of qualitative stuff, is that there's very few objective [screening] criteria that you're going to be able to use to tell you who's going to be able to perform interpreting or not. I'm one of those people that strongly objects to setting degree requirements for interpreting. An AA degree from a community college, is that necessarily going to say that you can be a competent sign language interpreter? I don't think so. There have been some studies of the relationship between degrees and interpreting performance, and there ain't none. You can give me all the college professors you want to take the federal exam or the state exam. Spanish college professors. And [with respect to] the relationship between being really competent monolingually and being able to interpret, you know, there's a lot of difference.

Q: My reason for asking was I know some states have a training requirement prior to taking the test. In those states, it seems that a lot of the people who might fail learn actually what it is to be an interpreter. And some of them realize that maybe they want to wait and get some more training.

A: Exactly. That's one real value in having a training component as a prerequisite to taking the exam. Even if the training doesn't do any good, the people might figure out during the training program, "Hey I can't do this." Also, just the fees, paying the fee and so on and so forth, these are deterrents to a whimsical decision of a non-competent person to take the exam. An interesting fact with respect to areas of specialized knowledge: your misjudgment of your ability is higher the more ignorant you are. The less you understand about something, the more likely you might be to say, "Oh, I can do this."

Q: I know in Washington State at least, there's a constant concern that there aren't enough certified court interpreters. And it appears that in the courts where they don't have enough, they just use uncertified court interpreters. So I'm wondering, what do we gain? I perceive this is going to happen in medical interpreting as well. We're not going to be able to certify enough people, so I'm afraid hospitals will just say, "Well, we'll just use uncertified people." So, why would you go get certified?

A: One of the things I'm a strong believer in is you've got to start somewhere. You've got to announce a standard, establish a standard. If you don't have it, you can't move toward professionalization. Many of you may be familiar with Holly Mikkelson's article on community interpreting, and all of the barriers to professionalization. You've got to announce a standard. And the political tension to lower the standard is great. By the way, [with regard to] RIDs lessons, I agree, almost everything Laurie Swabey said is right. All of our experience is the same. I agree and so I don't need to say anything more. There's a lot of evolution in the testing program. By the way [RID's] was one of the first testing programs I began to look at, remembering part of the reason that there were no standard tests in interpreting is that there's no professional organization, nothing comparable to RID. I studied or observed the NAD-RID wars. Oh, by the way, in interpreting (and I think this is really a lesson to take home), there is still to this day no strong, viable professional association on a national level in the interpreting community. What organization is there that would pretend to that? Well, ATA, yeah. That is a highly

evolved professional association, but translating and interpreting are not the same thing. There's something called NAJIT, and it aspires perhaps. It is the most influential and numerous association in the United States in interpreting. But what does its name stand for? National Association of Judiciary Interpreters and Translators. If I have one take-home sort of comment to make, and that's throughout my experience, I'm absolutely convinced that we do need [to discuss] in this country the notion of what interpreting is, and to have examinations and standards for being an interpreter, and then worry about the specializations later. I think all of the issues about specializing in a medical field, a legal field and so on are minuscule compared to the challenges associated with developing a professional cadre of interpreters. And that's never going to change until we can attack the whole employment environment and turn some things around. It's a negative vicious cycle between failure to appreciate and compensate interpreters, which goes hand in hand with the question, "What rational person would commit themselves to a career choice of being an interpreter when they get treated like crap. They get underpaid. You know, nobody values their services. And if you're in a language outside of Spanish, maybe in your city or in a court environment, you would get invited to work one day every month or something like that.

Lessons learned from interpreter certification initiatives at the state level and recommendations based on these experiences.

Pilot Certification Testing of Health Care Interpreters: Lessons Learned and Recommendations

Elizabeth Nguyen, California Healthcare Interpreters Association (CHIA)

I am here to represent CHIA as a board member and former co-chair of the CHIA Standards and Certification committee.

The California Healthcare Interpreters Association is a 501-C3 public charity dedicated to serving the public good. Our mission is health care interpreters and providers working together to overcome linguistic and cultural barriers to high quality care.

I was asked to share the findings and lessons learned from the joint pilot held in 2002-2003. This was a collaborative project between Massachusetts Medical Interpreters Association (MMIA) and California Health Care Interpreting Association (CHIA) to test the prototype of a certification exam created by the MMIA (now IMIA). The pilot was made possible through funding provided by the Office of Minority Health to the National Council on Interpreting in Health Care (NCIHC).

The purpose of the joint project was for NCIHC to support regional efforts, for the MMIA to obtain a larger sample size for its testing prototype, and for CHIA to learn lessons that will guide the recommendations for future certification.

CHIA's participation in the joint pilot was to assist the MMIA in the selection and training of the test administrators, proctors, and raters; to recruit and identify the test candidates; to support the coordination and preparation process; to assist in the implementation; and at the end of the testing, to discuss and provide some of the feedback. I would like to clarify the role of the CHIA Standards and Certification Committee (S&CC) during that time, and in relation to the pilot testing. At the time of the joint project, we had just developed and published the CHIA *California Standards of Practice for Health Care Interpreters* in September 2002, with funding from The California Endowment. While we were interested in gaining experience with the pilot testing, the S&C committee was not at all involved in the design of the testing, nor in determining the validity or reliability of that testing tool.

About the testing tool: the IMIA medical interpreter assessment certification tool has four modules. In module number one, there is a written test regarding medical terminology. In module two, there is another written test that has to do with standards related to ethical and cultural issues. Module three covers language conversion from Spanish to English and English to Spanish. And, of course, there is a role play that would test the testers on how they would perform in a certain situation.

We proceeded with selecting the administrators, and these are some of the basic characteristics identified in choosing them. They needed to have three or more years of experience as paid medical interpreters. They needed to be respected by others in their field. So we asked for references. They had to have experience training or

supervising interpreters. And of course, they had to have a commitment and a passion for the field and for the development of the profession. And we asked that they also engage in ongoing professional development.

Some specifics about the administrators in California: we had thirteen people. And the countries that they represented ranged from Argentina all the way to Spain. They are highly experienced interpreters. They have worked in their field at least ten years. They have diverse educational experience and background. They are aware of Spanish regional differences and they are quite sensitive to language issues among heritage speakers. We are in California where there are about 334 languages represented. And within the Spanish language alone, there are many regionalisms, so identifying the characteristics of these administrators was very important.

How did we recruit the participants? Our goal was to identify participants who represent three levels of proficiency, from “beginner” all the way to “advanced”. They had to fill out a questionnaire where we asked them about their employment history as medical interpreters. At that time, we also included people who just said, “Well I just volunteer as an interpreter.” We also asked about the type and the amount of specific training or education they had received. This is self-declared. Also, we asked them about the estimated number of interpreter encounters or the numbers of hours they had worked as interpreters. So it was a pretty basic questionnaire.

We conducted this pilot in partnership with Healthy House. Healthy House is a community-based organization, based in Merced, California, in the Central Valley. We had forty-six participants representing California. This is how we prepared the candidates for the test. We told them the essential readings that they had to have, such as the *California Standards for Health Care Interpreters*. Bear in mind that, at the time the pilot testing started, we had just published the Standards book. It was just September 2002 when we published it, and so we didn’t have the time to go out there and train people on the standards for certification, let alone expect them to understand all the sections in that book. We also asked them to read the working papers series that were published by the National Council on Interpreting in Health Care. So it was really not a lot of preparation.

We also *recommended* that they take forty hours of basic interpreter training, such as *Bridging the Gap* or *Connecting Worlds*. And you’ll notice that I underline the word recommended. It wasn’t as though we said the pre-requisite for participating is that you have to have had this training. This is what they needed to have in order to prepare for the test: the *Bridging the Gap Interpreter’s Handbook*, *The Interpreter’s Edge* by Holly Mikkelsen, the *Human Anatomy Coloring Book*, *Medical Terminology: A Short Course*, *Medical Standards of Practice*, and some medical dictionaries.

These were the lessons we learned regarding the test modules: We learned that the written test had its own value. The arguments for it are that the written modules measure basic knowledge. We asked people questions, and they could check-mark with pencil and paper. That could tell us their competency as far as entry level. It was noted in the report on national testing published by the National Council and written by Cindy Roat and Maria Paz Avery that the candidates who fail to demonstrate this basic knowledge should not be certified.

Now the argument against the written module is that, of course, written tests have no value in predicting the success of the overall language conversion. Written tests, as you know, cannot really determine how a person can perform and demonstrate her skills when she interprets. So again that was the lesson learned.

It was important of course to have language conversion, because that was a good predictor of how this person is going to do in the role play. And the role plays themselves were very, very good, as noted in the report. They were very authentic modes of assessment. But of course they were very costly to implement.

Test administration: Between Massachusetts and California, there were a lot of variations in the test administration. There were discrepancies and inconsistencies in that on the written test in Massachusetts, there was no time limit for modules number one and number two, the written tests covering standards and medical terminology. Yet, in California, because of miscommunication, somehow, we gave the participants a one-hour limit, which meant that most of the candidates could only complete module number one, which was the medical terminology. They were not able to complete module number two, which had to do with the standards of practice. And of course, again, I would like to repeat that at that time, standards of practice in California were barely known. We had just published them, and I doubt that a lot of the candidates in California knew much about the standards of practice that existed in Massachusetts.

The logistics for administering the language conversion were affected by opposite conditions. In Massachusetts, the test candidates came into a language lab and so the language conversion was proctored by administrators who controlled the process. In California, we did not have a language lab. The candidates came in. They had two tape recorders. On one tape, they listened to the prompts and then they answered with the other tape recorder. So they were the people who monitored their own answering, listening and all of that. So again, that was not quite consistent.

There were challenges in the administration of role plays. Again, there were too many discrepancies. The administrators were not consistent, allowing different pauses, for example. Inter-coder reliability was another issue. I believe that this was published in the final report. There was a very high variability in inter-coder agreement, and I think the possible contributing factors probably included failure to understand the coding system, not enough training, not enough detail in training, not enough explanation, lack of familiarity. All of these are factors that contribute to the high variability.

Just to put a human face on this whole process, I interviewed the raters myself. And this is what they said: "I think the pool of raters was very good, but because we had many different backgrounds, I think that some people were not as skillful as others in rating tests." "Some teachers approached the rating in a different way because they are more used to being consistent." Another comment here: "They told us how to rate the test, but they were not really explaining to us the rating process. I remember we had disagreements. There was a lot of disagreement regarding language equivalency or cultural issues, and so it was very ambiguous for most of the raters."

The implementation process suffered from lack of time and of course, lack of resources. Most of it was done by a lot of volunteer time, so that impacted the quality of the process. With regard to the testing instrument of course, we needed to be able to think of how to transfer the test over to other language pairs. I think we need to have a more stringent screening process, separating coders from administrators at the time of

their testing. The people who administered were the same as the people who coded. The biggest challenge was to convey to coders what represents correct interpretation for concepts that didn't have equivalency. This was one of the biggest challenges for the coders. We needed more funding. We needed people who would know how to do coding. For the participants, this is what I would like to emphasize: more reliable measures of participants' characteristics. To rely on participants to just declare how many training hours they've had is not an appropriate measure of the quality of training received. Nor can we rely on how many hours they've worked as interpreters to guarantee that they are competent either.

What has this joint project taught CHIA? To quote Don Schinske, our executive director, "CHIA does not have the organizational capacity to develop a certification test, set training standards, and promote certification for adoption by the state or, independently of the state, for dozens of health plans, medical groups, five hundred plus hospitals or eight hundred plus clinics." Consequently, our position is that CHIA does not endorse any certification process, but recognizes the major significance of the work done by others. We are willing and open to participate with all key stakeholders at either the regional or national level to explore the possibility of national certification.

Although CHIA, as an organization, does not have the resources to start a process for certification, we have identified some of the needs that must precede certification. We identified that it was important to have language proficiency assessment as a prerequisite for the test and for the training. People have to have standards of training, and then clear definitions for specific areas of training and knowledge. Then after the training, there must be some sort of post-training to see if the students are ready to interpret before they can even prepare themselves to take the test.

Because CHIA was not ready to undertake the certification process, the CHIA Standards and Certification committee disbanded and created the Education committee to start training and educating people. CHIA also proceeded to create an online registry of interpreters. It is like a directory with skill summary sections that allow the interpreters to go online and enter their contact information, and all information relevant to their work history and experience. At this point, there are no prerequisites -- such as years of experience or degrees -- to join the registry. However, it serves as an informed marketplace for healthcare interpreters who can compare their own skills with those listed by others.

Q: Building on this registry idea, is that currently up and alive?

A: Yes. Before I left, Don promised me that it will be online. But only members can access this at this point in time. So the requirement right now is that you have to be a current CHIA member in order to access that registry.

Maria-Paz Avery: Being that I helped developed the registry, I'm giving a presentation on this at the UMTIA conference on Saturday. And one of the things that Don empowered me to say, so I hope it's ok if I jump in here and say it, is that CHIA developed this with money from the California Endowment. CHIA is very interested, once it gets it up and running and has worked out all the bugs, to share both the technology, the program, all

of it with any organization that is interested. He suggests that you might want to wait a couple of months so that CHIA incurs the cost of working out the bugs.

Q: So does that mean that health care providers that want to access the list would also then need to become members? Are there organizational members or just individual members? Do you have to be a member to search for an interpreter?

A: Yes, at this point in time. We have members who are organizations. We have managers of language services. We have interpreters. We have agencies. We have free-lance interpreters. So if you are a member you can access the registry, to search, and to look for the people whom you want to hire. In creating the registry we were hoping for people to realize that this is where the field is at this point in time. The message was: "We cannot guarantee quality as much as you want, but if you looked down this list and into this registry, you will see that there are interpreters who have more experience than others or more training than others, and that's a good point to start with".

Q: I just wanted to ask a little bit about the rater training. Correct me if I'm wrong, but is it safe to say that one of the lessons that you learned from this process was, based on some of the raters' comments, that the rater training was not for enough time? I've read some of the comments here, "Not enough, perhaps the selection process wasn't the best." I believe it says twelve hours of training. I'm just wondering how it was established that twelve hours would be sufficient? Or was it just kind of estimated based on what material needed to be covered?

A--Maria-Paz: It was based on the fact that we had very little funding. What would the funding support? Obviously, we knew from the very beginning that that was not going to be enough. Definitely for California. It was even less than Massachusetts, because I went there for two days and that was all we could do. And then I left, and came back to Massachusetts. And in Massachusetts, we continued beyond the two days because we could do it. And that's one of the lessons I'll talk about.

Q: So can I just ask quickly then how many hours training you did have in Massachusetts?

A--Maria-Paz: Well I don't know exactly, because we had the two days of training. And in addition to that, we continued to meet over the course of time that we were rating, just talking about issues and so on. In Massachusetts, we had also been spending a lot more time across the state with many interpreters, talking about our standards of practice, talking about what it meant, etc. So it was a very different context.

A--Elizabeth: Right. I think what happened in California was that the state was so big; the volunteers came from all over the place. We didn't have enough time for them. They were not very familiar with the MMIA standards of practice. We were not able to be there personally for them to kind of guide them through the process.

Maria-Paz: Somebody asked about the registry. MMIA has had a national and now an international registry for several years. At first it was just name of interpreter, address, phone number and email. And we changed it about a year ago to really relay the qualifications, educational experience, interpreter training, affiliation with other associations, and a few more items. And now, about three months ago, we created the

possibility to upload the interpreter's resume onto that directory. So, I'd be happy to share what it looks like. But to the issue of being open or closed: many associations have an open registry where anybody can look into it. And what happens sometimes is that there are security issues with regards to the information that's on the website, and also scams that occur that specifically target interpreters. So we took it off last August -- July or August -- and now only corporate members can access that directory. So being public or not is an issue. Some organizations have chosen, for liability purposes, to take information off the public domain. So, I'd be happy to share the information of some of our lessons learned.

Q: Was there a scoring guide when you did your rater training?

A--Elizabeth: I am sure there was a scoring guide.

A--Maria-Paz: We actually developed the rubrics for certain areas. So that was the scoring guide.

Q: So it was more primary trait, like you were looking at different areas instead of just underlined items?

A: Yes. Well, it depended on which part of the test it was.

Q: The other thing is did you have parallel versions? So were you dealing with different versions of the test?

A: No, it was exactly the same.

Q: Were the raters scoring one person and then that was calibrated so that they all came up with the same rating?

A--Maria-Paz: We were not that sophisticated. Again, you know we were doing this on zero dollars basically.

Lessons Affirmed and Learned: The MMIA Medical Interpreting Assessment for Certification Pilot; Forum on Meaning and Accuracy in Meaning; Forums on Certification

Maria-Paz Beltran Avery, Massachusetts Medical Interpreters Association and Education Development Center, Inc.

Let me just say that it's really a pleasure to be here. Right now my mind is reeling so much that I don't even know where to start. Everything I have heard so far has made me want to start discussing the issues and not just listening. So, I'm going to work very hard to do the presentation I have prepared and not try to start a dialogue based on what others have already presented.

I also want to thank Elizabeth for actually doing my presentation. It's really interesting to hear another voice from a different position with a different perspective of the same project.

So, today, I will share a few things with you about the MMIA Medical Interpreting Assessment for Certification Pilot. I titled my presentation "Lessons Affirmed and Learned." As I thought about our experiences in the MMIA, and my own personal experiences with this pilot as well as with the pre-session on meaning that I did at the MMIA Conference two years ago ... [And to those of you who were there, I apologize. I have not written that session up, but I still have all the notes, and eventually I hope it will be written.] and recently with the forums on national certification which I first designed for a presentation at the Quality Health Care Conference in Seattle and that are now being conducted in other places across the country, I realized that we have begun to gather a lot of data on how the nation is thinking about certification. These experiences have affirmed lessons that we already knew, but they also generated new lessons for me.

I guess I was hoping, as I was thinking about lessons learned, that maybe we were already beyond those basic lessons. But unfortunately what these experiences did for me was to affirm a lot of the things that I think you've already heard from other people. But there were also other things that came up that really were lessons learned.

Let me start with a quick summary of the development of the MMIA prototype instrument - the MMIA Medical Interpreting Assessment for Certification (MIAC). The development of this prototype was an outgrowth of the work of the Standards Committee of the MMIA that had created and published the MMIA Standards of Practice in collaboration with EDC [Education Development Center, Inc.], which is where I work and where I directed a project to develop a college-level certificate program to prepare bilingual adults as medical interpreters. We started that project with Spanish speakers and Khmai speakers in 1992. At that time there was really very little in the field of health care interpreting, and so we were creating everything practically from scratch, even though we did work a lot with people who were ASL interpreters and educators. But we felt that the context of medical interpreting was very different from where ASL interpreting was, and, even at that time, we defined our approach to interpreting very differently. ASL interpreting, at that time, was still in the black box model and we were more into an interactive, culturally-based model. Out of this program, we developed the MMIA standards of practice using the DACUM process (a job analysis process) and based the development of the certification on these standards of practice.

We developed the prototype in Spanish because that's the biggest language group that we had, and most of us on the committee were Spanish speakers. We pre-piloted a version of the prototype in Massachusetts in January 2001, and then conducted a more formal pilot with a revised version and included CHIA in this pilot. Someone mentioned earlier a cost of a million dollars for development of a certification instrument. Our budget was zero. Eventually, we got, I think, twenty-five thousand from OMH through NCIHC to do the pilot that involved CHIA. That was like peanuts. Well, peanuts are good, but you need more.

We trained scorers for the pilot but I won't go through that because you've heard from Elizabeth [Nguyen] about the two-day training we did in both California and Massachusetts. The instrument itself was administered to thirty-seven people in Massachusetts and forty-six in California over two days. The first day we administered the written sections and the oral sentence conversion section. The second day we administered the role plays.

However, the final analysis included only forty-two tests. So, out of eighty-three people who took the test, we were able to use only forty-two of those tests for statistical analysis because there was a lot of incomplete data. The major reason for the amount of incomplete data had to do with the inconsistency of administration across the two sites. Also, because of inconsistencies in administration, especially on the written part of the test, we focused the analysis only on the sentence conversion part and the role plays. And even with the sentence conversion part, we had to eliminate some of the data. In Massachusetts we administered the sentence conversion section using a language lab, which was very controlled. For California, because they did not have access to a language lab, they used tape recorders to administer the sentence conversions. We had given instructions that once the audio tape was turned on, the test taker should have left it alone – that is, not stopped it at any time. The audio tapes were timed so that the test taker had a specific amount of time in which to provide the conversion. But we could tell, sometimes, from the recordings that there were people who stopped the audio tape. Where that was obvious, we threw those results out and used only those that seemed to have followed the instructions. That is how we ended up with only 42 analyzable tests out of 83.

One of the purposes of our pilot was to assess the validity of the prototype test. Does it measure what it's intended to measure? Was it reliable – here, we focused only on inter-rater reliability. Another purpose was to try out different methodologies. So we included different and more formats for testing the same thing than we would need in an actual exam. We wanted to see what formats worked best and what didn't work for different people. We also wanted to determine whether any of the modules could be used as screeners prior to administering the role plays.

As a committee, we spent a lot of time talking about why we would want to have certification. The purposes we came up with were: to determine basic entry-level proficiency; to provide a standard of quality to the consumers of the service; and to provide interpreters with an assessment of their proficiency.

We also spent a lot of time discussing principles of assessment, and those principles that we wanted to develop the test around. Whether we were successful or not, I'm still not sure. But these were the things that we were aiming for:

- We wanted clear and public content standards. We felt that we had those because we were basing the content of the instrument on the MMIA Standards of Practice.
- We wanted to have clear and public performance standards. Again, those of you who know the MMIA Standards know that it's not just a statement of a task or a function. It's also a description of what it looks like if you are able to do the task in a masterly way and what it looks if you lacked mastery.
- We wanted to use authentic assessment methodologies. We felt that the role play was at least one way of getting to that authenticity of the assessment.
- We wanted to address issues of equity, and for us, it meant that the resources to acquire the knowledge and skills should be available to the candidate prior to taking the test. So whether that's training or education, it's got to be in place. Or if it's about the methodology and formats used in the instrument, we wanted to make sure that the methodologies and formats were accessible to all and comparable across languages while accommodating cultural and linguistic differences. We didn't want the format and methodology of the test to get in the way of the candidate being able to demonstrate what they really were able to do and knew.
- And we wanted to be able to meet the criterion of consequential validity. We wanted the process and the instrument to be designed with a concern for the social consequences of the measurement. That's the hardest piece to do without a long-term process. Consequential validity refers to the accuracy of the decisions made on the basis of the instrument.

The instrument had four sections:

- Section one: knowledge of basic human anatomy and medical terminology vocabulary for which we used paper and pencil, diagrams, equivalencies, and matching terms with definitions.
- Section two: understanding of ethical and cultural issues. We felt strongly that cultural issues and ethical issues had to be part of the assessment process. We measured the cultural issues in two ways. We addressed them in the written section by using scenarios. But we also addressed them in one of the role plays, in which we posed a cultural dilemma. We wanted to see how the candidate would react to the dilemma. This part was scored separately and not as part of the accuracy score. "How did you do on that aspect of addressing the cultural issue?"
- The third section had to do with the ability to convert oral messages accurately and completely. This was done through the sentence conversion. "Can you go from L1 to L2? Can you go from L2 to L1?"
- Finally, the fourth section integrated all the knowledge and skills into the simulation, which is the role play. And that's the part that is costly. That's what's time consuming and labor intensive, and difficult to ensure consistency in the administration. That's what's the hardest to score.

So, in terms of scoring for accuracy in the role plays and also for the oral sentence conversions, we looked at mistakes, omissions and additions, but then we simply used a "mistake" score as the overall score.

What did we find? We found that the sentence conversion section did predict how well the candidate would do on the role plays. But what was interesting about this

finding was that it was only the Spanish to English conversion score that was the good predictor of role play performance; not the English to Spanish. We can talk about why that happened later.

In terms of overall predictive validity of the test score to on-the-job performance, we obviously had no available data to test for that. [With regard to] consequential validity, we also had no data to test for that.

With regard to reliability, we focused on inter-rater reliability. Each test was scored by two independent scorers. In general, our inter-coder reliability was pretty reasonable, but when we looked at the degree of agreement between specific pairs, there was great variability, which really says that some people were better raters than others. [With regard to] inter-coder reliability on the sentence conversion, sixty-two per cent of the coder pairs had inter-coder reliability of .80 or higher in the English to Spanish sentence conversion, and eighty-six per cent had inter-coder reliability of .80 or higher in the Spanish to English. Again it's interesting to see that there's a difference in the direction of the findings. A t-test showed that Massachusetts coders had significantly better inter-coder agreement than California coders. I've already talked about some of the reasons why that might have been.

How did the test takers perform when you look at who passed the sentence conversions from English to Spanish? Who passed the sentence conversions from Spanish to English, and who passed the role plays? What I want to draw your attention to here is that five people who passed the role plays failed the sentence conversion. What that said to us was that actually our sentence conversion module was much more difficult than the role play itself. And why was that? In the role play, we were looking not just at how accurate and complete the conversions were, but we were also looking at the auxiliary skills the candidates were using to compensate for maybe not having the highest level of language proficiency, particularly in English. But they were able to use auxiliary skills like asking for a pause, a repetition, an explanation in order to maintain accuracy and completeness. On the other hand, we also looked at the manner in which they used these skills – in other words, were they *skillful in using these skills*. Let me give you an example: In the pre-pilot, I was administering the role plays, and in one instance there was this candidate who kept interrupting after every five words or so, not even at logical places to stop. I was playing the role of the provider and I found myself getting so impatient I wanted to scream, really. I know that if I were a real provider, I would have been pissed, you know? This is not who I want helping me. So, we looked at that aspect - how skillfully was a candidate able to use the auxiliary skills in order to maintain accuracy and completeness?

What are the next steps for MMIA? What we really need at this point is to develop the blueprint for the instrument based on what we learned from the pilots. What we have is a prototype. So, when people call and say, "Can we use your test?" I say "No, you can't use our test, because the test is just a prototype." What we need is a blueprint that will allow us to create comparable forms, different forms of the test to be given at different points in time, but that maintain the same level of difficulty across the forms, that have the same number of scoreable items, that address lexical issues, that address idiomatic expressions, all of those kinds of things. So what we really need is a blueprint and as part of this blueprint, we need an item bank for each of the modules that we can draw on. We also need to test the blueprint, based on this prototype, with other languages and document the modifications that might be needed, but still again keeping

in mind consistency and comparability across forms. We also have to do a better job with training and screening the administrators as well as screening the process.

So what have I affirmed? Measuring the skill of spoken language interpreting is a complex endeavor. That's my "duh" affirmation. We're still at that stage. We're still trying to get that message across to our constituents across the country. Here's another one: excellent interpreters do not necessarily make excellent administrators or scorers. These functions require different skills. We, I know, should have done a better job of screening out some of the people that we had as administrators and scorers.

Many stakeholders in the field of health care interpreting do not understand the complexities of developing a certification process, especially one that has to address multiple languages, cultures, and ways of knowing and expressing that knowledge, let alone understanding how to develop a process that is valid and reliable across many administrations and versions. We in the field of health care interpreting have to come to a common understanding of what we mean by certification. We need to get away from just describing what people are currently doing that they call certification and say, "No. These are the requirements for good certification, and that's what we need to hold people to. If they don't meet these criteria, then it's not a valid certification." In addition, as we develop rigorous, valid and reliable assessments in a limited number of languages, we also need to develop other rigorous ways of measuring and acknowledging the competencies of interpreters of other languages.

There are a lot of questions about which we have little empirical data, I think. For example, do we know what the minimum level of language proficiency is in order to achieve accuracy and intelligibility? It's good to say that we need to have the highest level of language proficiency. But that is an unrealistic demand. We need to recognize that we're not going to get that all the time. When we developed the MMIA Standards of Practice, we did it very consciously knowing that we had to pay attention to the auxiliary skills that would allow an interpreter who did not have the *highest* level of proficiency, but who had an *adequate* level of language proficiency in both languages to be able to maintain accuracy and completeness.

How do we measure accuracy of meaning, and what does it mean to have accuracy of meaning? Do we have agreement on what we are measuring when we say that? What do we mean by equivalency in meaning? Can there be accuracy without completeness? How should we establish our cut scores? Is there a *good enough* score, and if so, how do we know it is good enough?

I want to share a couple of things that came up when we had that pre-conference on meaning at the 2006 MMIA Conference. Here are some of the themes that came out from what people said or wrote. There was agreement that meaning has many levels and aspects. The three big areas that everyone mentioned relate to the content. First, what is the semantic content of that message? Next, what are the context and the purpose of that message? Context and purpose contribute to meaning. And third, what is the intent of the speaker in conveying that message? That aspect of *intent* presents a lot of difficulties in terms of measurement. Is that something that we can really measure, the intent of the person? I know for myself, and I do a lot of work as a mediator, to get two people to understand the intent of what they said to each other takes a long, long time, if we ever get to it. So should we expect interpreters to be able to capture the intent of the meaning? I don't know.

What is equivalence of meaning in an interpretation? Here is what some people said: equivalence between the message in the source language and the target language. Did the message get through? Is it likely to be understood? Is there clarity of expression? Here's one quote: "*The message presented in one language elicits in the listener the same image, message, connotations in the mind of the speaker,*" and I put that in italics because I don't really know how we do that. Can we do that, really? Or rather is equivalence of meaning fidelity to the source message? But what does that mean anyway: does it mean equivalencies in the conversion, not in the understanding? Is it fidelity in the *conversion*, which means to me that the receiver of the conversion is able to respond to the message as if they had heard it in the source language? And if misunderstanding is part of that response, that's what happens.

What did I learn? If we are to move to national certification, there's a lot of work we need to do in arriving at a shared understanding of what we are measuring, how we are measuring what we say we want to measure and what the result means. And then we need to convey all of that clearly and explicitly to those who will be seeking certification and to those who will be availing themselves of the services of certified interpreters. In addition to developing an assessment for certification that is valid and reliable, we need to have agreement on the purpose of the certification among all the stakeholders. We can't assume that we share that same purpose. The certification process is not just about the assessment tool itself, and sometimes we forget that. I know I did. And we've now had many discussions to determine the requirements that will qualify a candidate to take the assessment. This, I think, is often the hidden but major point of disagreement among stakeholders based on different beliefs about things like formal education, status, literacy, etc. Certification is a political process and we shouldn't forget that.

What else have I learned? We need better training on ethics in general. Training on the profession's code of ethics appears to be inadequate based on responses to the test questions on ethics. In our test, we asked people not only to answer multiple-choice questions on ethical issues, but also to explain why they chose a particular response and how that response related to the code of ethics. If we had scored only on that basis, we would have failed almost everyone. Although they could check the right answer, they couldn't say why that response was the ethical thing to do based on the code of ethics.

We need ongoing, focused dialogue - this is what I recommend - ongoing, focused dialogue that builds a body of knowledge about what we know, what we don't know, and the mistakes we have made. We learned more from mistakes than we learned from the successes. We need ongoing, focused dialogue to develop a common understanding of what competent interpreting looks like at different levels.

We need ongoing research. I would strongly suggest that one of the things we should agree on is that all current certification efforts should include a data gathering and research component. Most of them do, but there should be an agreement that it should be publicly shared. Research questions should be developed collaboratively and contribute to a coherent body of knowledge, not just about measuring the skill of interpreting, but also about the theory of interpreting.

We need ongoing public awareness at every single level.

We need more rigorous training. Somebody noted the difference between training and education. I agree that there's a difference between training and education. We need training of trainers and educators. It's great that RID is doing that. But we need training not just on the "how-to's," the mechanics of interpreting. Maybe that's all that we can measure. However, I think for a really fully professional interpreter, someone who can do their job really well, they not only need to know the "how-to's," but they also need to know the "why-to's." I think that the "why-to's" are part of education. And when I say education, I don't mean just formal education. I think there are other ways of educating people. We need a common core of content, of knowledge and skills, and we need standards of excellence for training and educational programs.

Q: You mentioned the item bank, developing an item bank. You're not thinking of that in the context of the role play part, are you? And if so, what would you have in mind?

A: Well, I haven't thought it through, but you know it might be that we would have a sort of deposit of many different role plays that have been calibrated according to the blueprint. This brings up issues of security. I come from the world of K-12 education, and in statewide testing there is no way that any state would give the same test over and over again.

Q: You were just talking about developing parallel versions that have been equated according to your test specifications.

A: Yes, and you need a blueprint in order to do that.

Q: I'm concerned that the questions that you ask, which I don't hear being asked very often, get addressed during these days that we're here, specifically the ones related to how we decide on what's an equivalence and those kinds of things. Do we have any way to make sure those questions don't get lost?

A: Well, let's not forget that we're going to have a series of dialogues. And I think that we can begin to plan for those kinds of things. I don't think it's something we can answer right away quite honestly. I think the understanding comes from people talking about it rather than writing it down.

Q: I just wanted to get a clearer understanding of the sentence conversion part of the test. I see it's an oral, so do they hear a sentence and then...?

A: Yes. It actually is designed so that it starts with very simple, short, and not very dense utterances. And it progressively develops into more complex, denser, and longer utterances.

Q: So then it wasn't like a written statement that they would then convert, or like a sight translation?

A: No. This is all oral.

Q: It wasn't a sight translation?

A: No. That was a decision we made. We were talking about entry level. We were talking about the basic skill of spoken language interpreting. Further down the line we said, yes, we would -- we *should* have a module on sight translation. We should have other kinds of modules, but this was starting with the basics.

Q: As you went down your list of things that we need to do, I was reacting with, "Yeah, those are the kinds of questions that academics would demand." To what extent do you think those are important questions with respect to where we are on the practical utility of testing and legal defensibility? And I maybe know the answer in the way you sort of deflected a previous question.

A: I think in order for us to do a good job of actually measuring what we want to measure, we have to know what we mean by it. We, the people who are developing the tests, who are doing the scoring, etc, have to have a better understanding of what it means.

Q: Do you think that the published test specifications (let's say, of a court interpreter test like on the consortium website) are inadequate in terms of presenting the objectives of testing? You know, this is what we're trying to do?

A: I've looked at what you have on the website, and I think it's something that somebody could follow and do. But I'm not sure I really grasp the training that you do for people who will be scoring it. I agree with you. We get to a certain level of objectivity, but there is also a certain level still of subjectivity. But I think the more we're able to articulate what it means to each other, the better we'll be at coming to a consistent, objective measure.

Q: I mean for the federal test, these people are exposed to one full week of training that involves a lot of practice, automated feedback, and inter-rater reliability stuff. "Why did this get twenty percent..? Why did you diverge? I mean really intense, expensive stuff. And yet, in a real testing environment, and we've got statistics on this, there's roughly fifteen per cent of the scoring units that are not unanimous. And then among those, there's debate. We were able to figure out that fifteen per cent of the time, in a three-person situation, a minority view prevailed for the official score. I mean, isn't there inherently always going to be some renderings that your experts just aren't going to agree with? I mean no matter how much time you spend trying to articulate what conservation of meaning means.

A: I think articulation through the training is important, but then the people who are training the scorers need to understand what they're getting at. That's all I'm saying. I think we still need a lot of discussion, and you might have had it more than I think we have had it in the world of health care interpreting, because we're still very new at the measurement and testing piece.

Q: Other than the federal government, who can afford forty hours?

Q: I was just going back to one of your slides where you talked about equivalence. You know, you want to make the person who is hearing the utterance feel the way he would had that utterance been in his own language. And that is really the old Eugene Nida translation, kind of the fundamental goal of translation, and that's called dynamic equivalence. And really, that is what you're trying to do. And so I think it really is a good

model. But you have to kind of go back and deconstruct it to realize that really, that's the kind of equivalence that you're looking for. Because you're not only looking at the meaning of the utterance, but you're looking at how it was said so that you get all that intent and register and feeling.

A: Yeah. I totally agree with you.

Thank you.

[Editor's note: Dr. Avery was not able to include discussion of results of the NCIHC Open Forums on Healthcare Interpreter Certification in this presentation. For this, please see her paper, "Are We Ready for National Certification of Health Care Interpreters? A Summary of NCIHC Open Forums" (October 2007), which can be downloaded from the NCIHC website; go to <http://www.ncihc.org/> and click on Resources, Publications.]

State of Washington, DSHS Medical Interpreter Certification

Hungling Fu, Washington State DSHS

I'm from the Washington State Department of Social and Health Services. Here is a little background and a brief history of our program. I don't think I have time to go into great detail about the federal civil rights law and several state laws in our state, and there are five of them. They are all related to equal opportunity and non-discrimination, and some of them are specifically about bilingual testing and bilingual services, and interpreter services. Under these laws, within a few years time, we had a total of fifteen civil rights complaints and agreements. This covers the whole department and every administration within the department. Our department has six administrations, ranging from children's services to medical to juvenile rehab to all the civil services area, and all of these complaints [applied to] various administrations, so that covers the whole department.

Finally we come to the class action lawsuit against DSHS. This is the final piece I think that really pushed the department into the certification business, because we were under the gun so to speak: "You do it or see you in court." Then the agreement was signed in 1991, and after that this program was created. I came on board in March of 1991 to sort of start this process. So this is the issue here specific to limited English proficiency populations and service to them. It's a class action [suit]. You know this is a group, but it's not just this particular group that brought this issue against the department. It represents the whole array of [interests], [and while] the starter group is Spanish in a particular geographical region, the class action actually applies to the whole state and to other languages as well.

These are the responses from our department. There was a series of six studies to examine the problem we had at hand. And these studies were done statewide by different people and recommendations were made as to how and what the state can do to resolve the problem. Based on the recommendations by all of these study groups, a program called "minority affairs initiative" was established to address these specific problems. And finally the consent decree was signed as a result of the class action. The consent decree covers relevant laws and regulations, the corrective actions to be taken by the department, interpreter services, translation services, testing and certification of bilingual staff, contracted translators, contracted interpreters, and training of department staff as well as monitoring and compliance. These are the specific pieces that the legal service is monitoring on an ongoing basis.

The legal service I'm talking about is the firm that represented this class to bring the suit against the department, and in terms of testing and certification, they had been involved from the very beginning of the test development process. Based on the consent agreement, the department developed two policies to specifically address related issues. One is an admin policy. Another one is a personnel policy. Part of the admin policy specifically addressed the function of our program besides the services the department provides. And it was made very specific in these policies what we should do to make sure that we meet these Consent Decree Requirements. On the basis of these policies, the Language Interpreter Service and Translations office was created in 1991, with testing and certification as one of the functions of that office.

These are the steps we went through in developing the test. I'm not going to go through every one of the steps. I just want to point out that for the medical interpreter's test, we did involve professionals in the medical field throughout the test development process. Initially, the medical test development process involved monolingual medical professionals to review the list of medical terminology and all test materials. Later on, further in the process, we involved bilingual professionals, like bilingual medical doctors and nurses, and medical interpreter service coordinators, DSHS program managers, and so on.

Another thing I want to point out here is the rewriting of the test. I know that there are some concerns about [the fact that] the test was developed in Spanish and simply translated into other languages. That's not quite accurate. First, the translation issue was most likely related to our internal employee test, which had a different focus on the balance of two languages in the test, due to how the department of personnel tested potential state workers at that time. Secondly, if you look at our medical test, there is not much text that needed to be translated. What are in the target language are mostly single terms instead of solid text, which most languages, if not all, presumably have equivalent or similar expressions. Besides, the test was actually developed in English and adapted into different languages. There are ten questions on sentence completion that are in different foreign languages. It is true that the foreign language part was done by qualified translators. But the instruction to them was they recreated an equivalent version in their respective foreign languages based on the content, the style, the meaning, and the complexity level of the English original, instead of direct translation. When the foreign language part was reviewed and edited by a second professional, they were not given the English original so that their editing was not interfered by the English original. Instruction was given to the editors that the final product should read like original writing in that foreign language. But if anyone believes that the test was written in Spanish to begin with, then I think that I can take comfort from that, because the person who did the rewriting into the Spanish language had done a very good job. As a matter of fact, we have had very few complaints in that regard. When the same test was implemented in different languages, we had to concern ourselves with the question of how to establish and maintain a consistent standard across languages instead of just for one or two languages. That's the struggle that we had from the very beginning. And it's not just the instrument itself, but also the process of test grading. Test administration is another component that has to be consistent across all languages. For a test process with human factors involved, such as manual grading, consistency is directly related to its integrity, especially for a government agency like ours, it was, is, and always will be, constantly under all kinds of scrutiny. We don't want to have another major complaint or lawsuit against the process itself, [asking] "How come you treat the Spanish speaking population differently than the Vietnamese population? How come we are less important, because we have a lower quality group of interpreters," that kind of thing. So we were very aware of and paid special attention to that from the get-go.

Let me use the example of the Chinese language to address the, quote, "translation", unquote, concern here. Like I said, the reviewing of the prototype was [done] by another group of language specialists. And the original English was not provided to them when they did their review, or editing if you will. I happen to speak Chinese so I reviewed the final product of the Chinese test myself. I think they did a very good job. You can't even see any traits of translation in there. And the only thing we have heard concerns about from test candidates is mostly [regarding specific] words. Occasionally we might hear comments like, "this is not how I say this word in my

language,” rather than “this is not how I [convey that message] in my language.” There’s a difference there.

Let’s move on. Assessment by bilingual [raters].... test validation... This is what we got out of item analysis. You can see we ran this analysis on different dates. We are happy with the results there. Test reliability was also analyzed. And these are the results. Now for the oral test, we did two kinds of reliability study. One is the “test re-test” analysis. Because we’ve been doing this for so many years now, we have a lot of data to study. And we also had inter-rater reliability analysis. As these statistics show here, the test appears to be highly reliable.

This is interesting. I think it was a couple of years ago that our department switched the way language service was provided. Because of how the delivery system was set up, it greatly reduced the availability of interpreters serving our clients. There were a lot of factors that went into play in this. The reason why we switched the delivery system was mainly because of budgetary constraints. The legislature said, “You are spending too much. You might need to cut interpreter service”. And Medical Assistance, one of the administrations within our department responded, “Well maybe we can provide the service differently to save money. We want to keep the federal matching fund. Why don’t we reinvent this service delivery system so that we can have more cost-effective services?” And based on their numbers, they saved millions of dollars. That’s what they said, at least. But because of that, it has made this delivery system so cumbersome. And for the interpreters it is really bad. That’s what I can tell. Before, for example, we paid interpreters like thirty-eight dollars to thirty-nine dollars an hour. Now in this new system, you get two layers of service brokering in between the end user and the interpreter. Layer one is what they call brokers. They take orders from medical doctors and hospitals; they then contact those language agencies which contract with individual interpreters. We have around fifteen to twenty language agencies, I think, statewide. They contract with the brokers, not with the department. And the agencies dispatch interpreters to specific job assignments. Both agencies here scrape some dollars out of it, so what goes down to the interpreter is twenty to twenty-two dollars an hour. So now the interpreter says, “Forget it, it’s not worth my effort. I’m gone. I’m going to go do something else.”

Then came L & I, Labor and Industries. There’s a worker’s compensation program. They require their interpreters to be certified by us. They’re paying, depending on the language, forty to forty-five dollars per hour directly to the interpreters. Now you see what happens. Everybody goes to L & I. And we are left with inadequate resources. So that’s why at MA they were [complaining], “Well, we don’t have enough interpreters.” Consequently, they had all kinds of inquiries into the testing and certification program. “Oh you have to lower this, you have to lower that. Otherwise we won’t have enough resources. We need more interpreters”, even though we have certified thousands of interpreters, literally. I think we have over 4000 certified medical interpreters alone, starting from 1996.

They would ask, “Why do you have two sections in the oral test? One is the sight translation. Another one is the consecutive interpretation. Why don’t we just get rid of the sight translation so that we can have a higher passing rate?” even though our passing rate has been always around 37%-38% ever since 1996 and across languages. That’s when they talked about revamping or overhauling the process. They just wanted to take this apart, to take out the sight translation section of the oral test. That’s when I took a

sample out of my archives and did a study. The result we got showed that the sight translation and the consecutive interpretation are very strongly correlated. As shown in this scatter plot, you can see a few obvious outliers there somewhere, but overall they fall very close to each other. So after I presented these results they agreed that: “OK, let’s stay with that. Let’s stay with the format we have right now.”

Q: So, you kept it? You kept the sight translation?

A: We kept the sight translation and the consecutive interpretation together.

Q: Despite the high correlations?

A: Yes, because interpreters need both skills, we kept both.

Q: Well, some people who may not understand would tell you that because there’s a high correlation, you would want to take one of them out. However, we know that from an interpreting perspective, even though there’s a high correlation, they’re still measuring two very different psycho-linguistic and linguistic kinds of tasks, so that was very good.

OK. That’s the written test. That’s what it covers. The oral test: what are we looking for here? Test scoring.... Rater training. What we did at the beginning, is that we had a large group of raters. I think from the beginning, we had I believe sixteen to eighteen people for the Spanish language alone to go through rater training. What we did was [have] everyone come into our office, to work together as a group. I think it took about a week, and we had a large batch of tests there all ready to be graded. So, these people, during the training, worked together in pairs. Eventually, we retained three of them. And we still have these three people working since 1992. So they are very consistent. And even after the initial training, these three people sometimes work [together] on their own time. They use their own time...to get together and grade some tests together, to make sure that they keep the consistency. So the inter-rater reliability is pretty high. It is .95 or somewhere around there. We are pretty happy about that.

Of course availability is also a concern to us, and that is one of the reasons why we cannot yet enforce a policy of continuing education [or] recertification. If we could not have enough interpreters because of the requirements imposed, [then we would be in a mess.] Interpreter availability concern will come first from the programs. And then you have these clients out there who don’t get the service, or have interpreters, we will then really have a problem.

Panelist: We should be clear though that a provider that’s responsible for the interpreter can get an uncertified interpreter. They just can’t bill the state for it.

A: Yes.

Panelist: They can’t get them through the broker.

A: That’s true.

Panelist: So what's happening is that places could be getting interpreters for their patients, but they don't want to pay for it, so then they get no interpreter.

A: We don't have any more complaints or anything against the department regarding the quality of interpreter services. That's a good thing, I think. Here [are] the key considerations. I don't think I have time to go through these actually, but it's in the packet, if you have any questions or if you want to talk about it, we can probably talk more about these in the breakout sessions if we need to. These are the test development considerations, test administration, test scoring, and pre-test requirements.

Q: What are the pre-test requirements?

A: There are no restrictions on the eligibility for taking the test, such as formal education requirements. For example, they don't have to have a B.A. or A.A. degree in order to take the test. Again, we are a state agency, and discrimination concerns are always in the back of our minds in regards to [people who claim], "I can [interpret]. Why don't you give me the opportunity to try to show you I can do that?" We provide pre-test materials. What we have is a package that includes a pretty large size vocabulary that people can study. But most importantly is the format of the test, written and oral. There's an audio CD we send out to test candidates. What we are showing them are that these are the kinds of questions we are going to ask on the written test, and this is how the oral test is going to be done, so they get familiar with the process itself and don't get caught by surprise when they take the test.

Q: Did I hear correctly that you said that you were able to certify 4000 interpreters in just ten or eleven years?

A: Yes. We are talking about all languages, not just one language or the certified languages. Every month I think we have around fifty or sixty people certified in all languages.

Q: Just being a neighbor in Oregon, I know that a lot of Oregon folks have gone to the state of Washington to take the exam as well. And that may be true of Idaho also, so maybe those numbers include folks outside of Washington as well.

A: Initially, I think we had people from all over the country that called. Our phones were ringing like crazy, "Can we take your test?" We did test them at the beginning. But because of resource constraints and these people are from out of state, they can't serve our clients, so we decided not to do that anymore. We still do test people from Oregon on the border and people from Idaho on the other border. Those people can take the test; they can come across the border to serve our state.

Q: What's the cost of the exam and who covers the cost?

A: The test fee you mean?

Q: Yeah, so is the person who is taking the test paying for the test that they're taking or is...?

A: The written test is [still] thirty dollars. The oral test is forty-five. Seventy-five total if you take both. It only covers test administration expenses such as room rental, test proctoring cost, and test grading cost. We have contracts with some test proctors, the same individuals. Again, consistency is important. We kept the same [proctors] as long as possible, five of them. Three [are located] in Western Washington, and two in Eastern Washington, [working as] two separate teams.

Q: How much is it costing you per test to offer the test? So the person who is taking the test is spending seventy-five dollars. How much money is the state of Washington spending on each exam?

A: OK. As I said, we have several regular staff on state payroll. I am one of them, and we have another two people. One is working full time doing test scheduling, and another person helps in between. But the test fee (seventy-five dollars a pop) is not quite enough to cover the cost, so the state will just chip in and [cover] whatever is left over.

Q: What's your pass rate?

A: The pass rate is thirty-eight percent, around thirty-seven or thirty-eight percent.

Q: For a first time or across the board?

A: Across the board, ever since 1996. It fluctuates a little bit from time to time, but it's around thirty-eight percent for all languages.

Q: On your last slide, you had a point saying "Coordination Among User Agencies" that you didn't get to speak to. Just quickly, what do you mean?

A: In our state, our department is the only agency giving the certification testing. But agencies like L & I, (Labor and Industries), Licensing, and Corrections are all requiring their interpreters to be certified by [us], because they don't want to spend the money to reinvent the wheel and do something [for which] they don't have the resources. These agencies are fighting for resources, but they're paying differently. Like I said, L & I is paying forty to forty-five dollars an hour to interpreters. For us, interpreters get twenty. You fight for the same resources, and then we are left either with inadequate resources or inferior quality for that service. So coordination among these agencies would help to best utilize available resources.

Lessons learned from corporate and non-profit certification exams

Language Line Services' Medical Certification Program [presentation not available]

Janet Erickson-Johnson, T & I Certification Manager and Director of Testing,
Language Line Services

A Proprietary Certification Program for Telephone Interpreters: Development and Insights

Frances A. Butler, Ph.D., Language Testing Specialist
Consultant to *NetworkOmni Multilingual Communications*

Good morning everybody. I'm very pleased to be here, and I wanted to thank Bruce for inviting me and asking me to participate in this interesting and important discussion that is going to be taking place over the next two days. As you know, I'm an independent consultant in language testing, and what I'm going to be reporting on this morning is a project that I was involved in with NetworkOmni Multilingual Communications on the development of a proprietary certification program for telephone interpreters. I've added to the title of the presentation *Development and Insights*, and following the request of Veronica and Bruce, I'm going to move fairly quickly through the details of the program and spend a little bit more time on insights from the work, challenges that I see as a result of that work, and the current effort that we're here to address.

First, I plan to talk about the background and rationale for our approach and then steps followed in the initial program development, then ongoing reassessment and periodic updates, insights on the application of testing principles in a new assessment environment, that is working with interpreters, and then issues and challenges in the current effort. Some background and approach: we began by putting together an interdisciplinary team of experts to work with NetworkOmni staff. Details about the program, who the experts are and so forth, are covered in Cindy's report for the California Endowment. So, for more specific information, I refer you to that document. There is also, I believe, a copy of an article in your binder that was in the *ATA Chronicle* that is about early stages of the development of the work at NetworkOmni. It is under Sawyer, Butler, Turner and Stone. So those two articles will provide more specificity about the program that I'm talking about today. We had an interdisciplinary team of experts that was composed of language testers, psychometricians, people who had expertise in cultural competence as well as interpreter training. The approach that we took was building content from an empirical base. This was very important to us and important for helping to establish the validity of the process. I'll talk a little bit more about that empirical base in just a minute. But from the empirical base, decisions were made about the makeup of the program. The program is modular, and it has a multi-skills assessment. By modular, I mean there are different pieces of the program that focus on training and testing, monitoring and so forth. Within the various phases of the program or

modules of the program, we have a multiple-skills assessment that focuses on linguistic skills as well as interpreting skills.

The empirical base for the content, to go back to that for just a second, was important because for us it reflected the types of industries that NetworkOmni was working with and also the ranges of interpreting scenarios within and across those industries. I'll talk in a moment about how that was important feeding into the training and testing phases of the program. The steps followed in the development of the program were the following. We began with a needs analysis, and by this I mean we began by reviewing the existing processes. We went into this effort not starting from scratch because NetworkOmni already had a training and testing program in place, but what they were interested in was having us come in and help evaluate the program, refine it and develop it into a certification program for them, including an internal certification exam for their telephone interpreters.

So we reviewed the existing processes. We documented the client base in terms of the types of scenarios that interpreters were being faced with in day-to-day interpreting situations, and the documentation of these scenarios allowed us to articulate the types of skills that interpreters were being called upon to demonstrate, and also then what we needed to test. As part of the needs analysis, we reviewed the client calls across industries, and this was very critical for establishing these patterns of calls across industries so that we could make decisions for selection purposes for training and testing. And the analysis of these calls, as I mentioned before, led us to the identification of scenarios. And these scenarios then eventually fed into test specifications and the design for the training program and for the testing program. So once we had gathered this information, we drafted a framework for test development, and for the development of all the materials in fact. And the drafted framework led to the various stages. We laid out those stages—screening, training, testing and monitoring, and in some instances these stages became iterative, in a way similar to what Janet was just talking about with Language Line. If at certain phases the interpreters who were tested did not pass the test, there was an opportunity for them to cycle back and receive additional training, additional monitoring and have an opportunity to be tested again at that particular stage before moving forward.

We used the information that we had gathered from the empirical work to develop and refine the materials and the test. We did this, as I mentioned, through the development of test specifications and then actually drafting items and beginning the piloting process, the tryout process, and revisions, again iterative. Those of you who have worked extensively in test development know that, as Janet was saying a minute ago, this is a process that requires a lot of going back and refining, trying things out again to get things to be the way you'd like them to be.

So once we felt confident about the quality of the procedures, we conducted a validation study with clients. We went out to a number of the different clients from different industries. Through the piloting process we had recorded a number of interpreters actually taking the test. We played these recordings for clients and asked them a couple of questions. First of all, we asked, in terms of the content of the test and the types of interpreting tasks that interpreters were being asked to do, "If interpreters could handle the material on the test, would you feel confident that they could handle most of the types of calls and so forth within your company?" That was one type of information we got back from the clients.

Next, the performances that we showed clients were different ranges of proficiency in terms of the effectiveness of the interpreters who had taken the test. We wanted to get a sense from them of the minimal level of proficiency that they would consider acceptable, that they would feel confident that the interpreters could handle everything in a high quality way for their personnel. So that was the validation study that we carried out with clients.

One of the things that we felt good about was that with the validation study we did not have to go back and make a lot of changes. Most of the clients were very positively disposed towards what we had produced in terms of the tasks. I credit that with our having worked from an empirical base. Before we actually did the test development, we had a solid feel for the kinds of situations in which interpreters were going to be interpreting. So going from that base meant that we saved a lot of time later when we could have made mistakes if we hadn't had that solid background to build on. Working from an empirical base, I think, is always very important in a test development situation. Once we did the validity study with the clients, we then moved on to implementation of the internal program.

How a program is implemented is really important. It's something for us to think about in the current situation. With the NetworkOmni project, we didn't just implement the program and have it stop there. An important feature of the NetworkOmni program is the ongoing reassessment and updates to the program. These steps are critical for maintaining the validity of the whole process. I think it's important to remember that validity isn't a one-shot deal, but that it's an ongoing process that should continue for the life of any program. It's a part of keeping what you do current. So, on a periodic basis, NetworkOmni takes into account clients' changing needs. That could mean new clients with new needs coming in and being part of the program. It could also mean that there are internal changes going on in client companies, and that they may have different kinds of needs in terms of the types of interpreters that they need or the kinds of calls being handled and so forth.

Another element to think about is new approaches to training and testing. As we become more sophisticated in the techniques that we're able to use in testing and training, we want to be sure that those new techniques are implemented in an ongoing program. Yesterday someone mentioned the fact that changing technology has a lot to do with how we're able to deliver testing and training. So changing technology is an important piece of this too. And then as all these different changes are happening, it might mean the need to bring in additional outside expertise to help work on these various issues. So I think the main point here is that constantly monitoring the quality of your program means that you need to stay in close touch with your clients or with whoever the program is serving. As changes are going on, on a day-to-day basis, you need to figure out a mechanism for monitoring changes and getting the important ones implemented in whatever the program happens to be.

So now what I want to do is move on to the insights that I've had as a result of being able to work in this new testing environment. As a language tester and through the years I worked at UCLA I had an opportunity to work in many different language testing environments in this country and overseas, but until I came to NetworkOmni, I had never worked with interpreters. For me, it was an amazing opportunity and a challenge. It was extremely interesting in so many ways. One of the first things that I became excited to

learn about was the complexities of interpreting beyond language proficiency. Clearly language proficiency in whatever the two languages are is a very important part of interpreting, and language proficiency must be assessed. But clearly, and this has come up in discussions already before now, being able to actually tap into interpreting ability beyond language testing is an interesting challenge. And how you define interpretation skills in the various milieus in which interpreters are functioning is part of what ultimately goes into the test specifications for testing interpreters.

A second issue for me was the special context of telephone interpreting. From one interpreting environment to another, you're going to have different situations that have to be addressed in different ways in terms of evaluation. Working on the NetworkOmni project really reinforced for me the need for an empirical base. I feel very strongly that an empirical base is critical for helping to establish the validity programs and assessments within those programs. And there may be instances where you need to be able to show why you're testing in a certain way. The empirical information that you have provides the data that you can show the outside world, in terms of why the approach you're taking is the approach that you are taking. Also it will help you explain why you aren't doing something else. Another thing that I learned was the need for client education. It became very clear in a lot of instances that clients did not know how to work with interpreters. And so training clients to use interpreters effectively is very important, and I think actually somebody else mentioned that yesterday. I don't remember who it was.

Finally the issues and challenges I see in the current effort. One is the need for an articulation of what certification means in this context. We really need to be able to spell out what it is we're talking about when we say a person is certified. We had a little bit of discussion of that yesterday, and I think that in one of the smaller groups yesterday afternoon there was further discussion of that. But we need to be able to say what certification means and also what it doesn't mean.

Then, there's a need for clear specification of the content domains and sub-domains. I know there's been some discussion of whether we're talking about general community interpreting, or whether we're going to be talking about medical/health care interpreting. Whatever the decision is, in terms of the national certification, the content domain and the sub-domains have to be very clearly specified. And this is going to, I think, help feed into the articulation of what the certification means. We were talking a little bit in the testing group yesterday afternoon about this notion. Let's say we decided to focus on health care and medical interpreting. That would be a narrower focus than community interpreting, but still a very broad focus. And there is the question of how you make decisions for training and assessment. Everything cannot be covered or tested, so selection priorities must be set. Setting those priorities is one of the things that we need to do.

In closing, I just want to reiterate something I said a minute ago, and that is we always need to remember that validation is a process and that it's critically important to document every stage of the process. The meeting that we're attending here today is a step in a process. And I would say that depending on where the focus is in terms of community or health care interpreting, if it turns out that the standards that have already been developed are part of this process, then it's important to document how those standards were developed. In addition, there should be outside validation of those

standards if they are going to be critical in terms of informing the decisions that are made.

Q: I just wanted to ask about when you say multiple skills assessment. What did you mean exactly by that?

A: I meant that in the scenarios, in the testing, we were tapping both linguistic elements and interpreting skills. I didn't go into a lot of detail, in terms of diagram and sort of the flow of how the system is set up, but first there's a screening process, where language proficiency alone is initially tested for the two languages. Then as we move into different stages of training, testing and monitoring, at the later stages, language proficiency or linguistic skills continues to be part of the evaluation process at every level.

Q: You talked about a customer satisfaction kind of study, and I understand that. But your other study, where you asked clients what would be the minimal level of proficiency they would be comfortable with, and you played I guess different levels of interpretation, could I get some more details on that?

A: I'm not sure I can. Jean is here. I should say Jean was part of the team. Do you want to respond?

A (Jean Turner): Sure. As I recall, and I believe I recall accurately, there were three major client groups that were identified, and then we met with three clients within each one of those client groups. And there were two of us. Sometimes it was David Sawyer and I, or Frances and David, or some combination of the three of us who visited each one of these clients. And typically the people that we met with included the supervisor or the customer service representatives for the client, and a number of people who were customer service representatives who were bilingual. Because there was some question of how do you know whether the interpreter's doing a good job unless you know the languages? And if you need an interpreter, you don't know the languages. So how do you figure that out? So that was why we specifically requested there would be people that would know the languages as well, so that this process of listening to calls they could respond to something other than just whether their voices were nice.

Q: Were they interpreters or just some kind of bilingual...?

A (Jean): Two of the groups that I participated in, client groups, there were one or two or sometimes three people who actually served as interpreters themselves for the companies.

Q: So did you play failing performances?

A (Jean): I wouldn't say that they were failing, but that they were poor, weak.

A: And I'm not sure. I may have misspoken when I said we asked them to identify a level. We didn't use the term level. I guess maybe in our minds, we were trying to see it. The tapes that we played showed a range of ability or effectiveness or whatever terminology you want to use.

Q: But were any of them failing?

A: No.

A (Jean): We didn't think it would be productive or good P.R. to write a failing list because in fact people who failed at the initial stages of the program are not working for the company because they demonstrated that even with training and support, they couldn't handle the work. So, no.

Q: It just occurred to me. I'm sorry I didn't ask. Is this a medical certification program or just general or all different types?

A: It's a general certification program to begin with, and there's a module that focuses on medical.

Q: So, there's a separate test for medical and a separate one for another module?

A: It was constructed that way so that over time, as the client base might change, different areas could be added. But the initial approach is that we wanted to tap first basic interpreting skills that could cut across the different industries. And actually, I don't think I mentioned earlier, but as part of the empirical work, when we were looking at the types of calls that were coming in from across industries, we were looking at four patterns or types of calls and types of scenarios and so forth that would address the general interpreting skills.

Q: I was just going to ask you what domains or competencies actually were tested or are tested for certification? You mentioned language proficiency, but in terms of interpreting, do you test auxiliary interpreting skills? I'm just trying to understand what it is that you test.

A: General interpreting skills, abilities to transfer from one language to the other.

Q: So, consecutive interpreting skills?

A: Yes.

Q: Any other skills? Like sight translation?

A: No translation, no.

Q: So you said it's a basic skills test, and you have modules. But you're testing company bilingual employees? So different positions, different jobs have different modules that you test?

A: Well, there's the potential there for developing modules. But a lot of the work is customer service across a range of industries, so that's the general test that the program covers, the training and testing. But then it also offers the opportunity for more specific modules, and health care was one of the areas where a different focus was needed. So in addition to the general customer service type of material, health care was added as a separate module. And it would be possible to add other modules if the type of work

became more specific than what had already been identified from the general interpreting data.

Assessing Medical Interpreter Proficiency: The Medical Interpreter Competency Examination (MICE) [presentation not available]

Roseann González, Professor and Director, National Center for Interpretation Testing, Research and Policy, University of Arizona

Implementation and Validation of Language Proficiency and Interpreter Readiness Tests

Jean Turner, The California Endowment / Monterey Institute of International Studies

I am at the Monterey Institute, and the work that I have done on the test that I'm going to be talking about has been as a consultant to projects that were funded by The California Endowment. Those projects were funded through several different entities, so we weren't really sure how to characterize where these tests came from; therefore we decided to give the credit to the funding agency, which was The California Endowment. I am suggesting that we refer to them as the California Endowment Tests, because there are a lot of AKAs or aliases.

I'm going to describe these tests briefly, because there is actually a whole suite of tests. I'd like to note that some people refer to suites of tests as test batteries, and I think that that is an appropriate word for a suite of tests. I think it's something we should consider as we're thinking about a certification effort. Is it going to be a test battery, or is it going to be a single test?

So this test program that I'm going to be talking about includes four tests in several different language combinations. There are language proficiency tests in Spanish, Cantonese and Hmong, and there's also an English-language proficiency test. There's something that's called an Initial Interpreter Readiness Test and something that's called a Final Interpreter Readiness Test. I'm sharing this terminology with you because this is the terminology being used in various reports that are now being written about the validation work on the tests. I think it is helpful to begin to think of the tests by these names.

First, I'll describe briefly the language proficiency tests. The Spanish, Cantonese and Hmong language proficiency tests were designed and developed by Claudia Angelelli, which is why some people refer to them as Claudia's tests. They were developed with funding by the California Endowment for the Connecting Worlds Partnership or Consortium. There was also an English language proficiency test, which was developed by my colleague Renee Jourdenais and me. It's based on the design specifications for the other language proficiency tests, but it's a little bit different, because of the requirements for English being a little bit different than the requirements for some of the other languages, specifically Hmong. There is not a reading or writing component to the Hmong language proficiency test. The original purpose of these tests, since they were developed for the Connecting Worlds Consortium, was to determine whether the individuals who presented themselves for this training had the language proficiency in both their target language and English to engage in training.

The specific features of these tests include that each one of them has six tasks; those tasks are designed to measure listening and in some cases reading and speaking. The test content is delivered through audiotape or CD. There's a very small writing component on the Spanish, Cantonese and English language proficiency tests. There is no writing and there is no reading on the Hmong language proficiency test. I've listed a couple of example tasks, not all of them, because there's not time for that. I don't know that we really need to go into that depth anyway, but take as an example the reading task. Examinees might read a label on a medicine bottle and then answer

comprehension questions. By the way, all of the responses are oral and are recorded for later scoring. An example of a listening task might entail listening to a patient's description of a medical condition and answering comprehension questions regarding that. Another possible example is a speaking task where one listens to a health care provider describe a medical procedure and then, in the same language, simply explain that to a patient who may not have understood the register. It's a thirty-five minute administration time. There's scoring by two trained raters.

The Interpreter Readiness tests were developed in three language combinations: Spanish-English, Cantonese-English, and Hmong-English. Again, they were designed by Claudia, so these too are sometimes referred to as Claudia's tests or the Connecting Worlds tests, or the *Hablamos Juntos* tests, which I will get to in a few moments, or LESA, or the California Endowment tests. As I indicated previously, I'll refer to the suite of tests as the California Endowment tests. The original purposes of the Initial Interpreter's Readiness was as a screening or a pre-test prior to interpreter training, with scores to be used to determine whether someone had the requisite interpreting skills for engaging in the medical interpreting training that was offered. The final Interpreter Readiness test was designed to be given following training, to measure whether people had the requisite language and interpreting skills to function specifically as a community medical interpreter.

Here are some of the features of these tests. I'm going to call them the IR tests because I only have twenty minutes, and Initial Interpreter Readiness and Final Interpreter Readiness is going to take me a long time to say every time I have to say it. So both IR tests have four short video-taped segments of simulated interactions between a health care professional and a patient. The patient is Spanish-speaking, or Cantonese-speaking or Hmong-speaking, and the health care professional is English-speaking. The examinee takes the role of the interpreter viewing the role play or the scenarios. There are pauses for that person to do interpreting. The responses are recorded for scoring, which is done by two trained raters. Scoring is done by people who are proficient in both English and Spanish, and there is a scoring rubric that presents the script for the scenarios, organized by turns. And then there are critical components of the turns that are scored separately, distinctly for accuracy: one point for accurate, no point for inaccurate, as judged by the trained raters. Not all turns are scored, and not all of the information in each turn is scored.

Here are some of the initial IR tasks. By the way, there are four scenarios, but it's the same patient engaging in several different tasks, things that patients do. So first, a patient makes and verifies an appointment, changes an appointment, has the appointment with a doctor, and then makes an urgent care appointment request and interacts with several health care professionals in that. That one is pediatric, by the way. In the final IR tasks, the patient makes an appointment and sees a specialist. There is sight translation of a consent form, and then an admission process in the third scenario, and then the fourth is a follow-up appointment. Perhaps interestingly to some of you, the sight translation in the trials that were done of the Hmong-English version of the test was not performed by any of the candidates.

Those are the tests. Those were developed with funding by the California Endowment and by Claudia. I was a consultant on that project. My involvement in the first and following phase of the implementation and validation work was as project manager for that effort, separate and distinct from the test development process. The

goals of this first phase of the validation work that was completed in 2006 were to support the implementation of the test. The tests had been developed, but they were not in use.

We presented the tests to the five organizations that had contracted them, and trained people at the sites in how to administer the tests. We identified and trained a pool of raters at each one of the sites, or a rater at each one of the sites. We also determined rater reliability, and we had planned to also investigate the impact of training on test performance and the impact of the amount of experience on test performance. This is the information about the reliability in general for the tests that were given. We had thought that we would be able to do this battery of testing with twenty to twenty-five people at each of the sites. In fact, we ended up with a very different configuration of people. For example, there was a total of fifty-one people who did the English test, the English language proficiency test, and only six people who did the English-Cantonese Final Interpreter Readiness Test. The rater reliability ranged from 0.55 for the Cantonese Final IR, to 0.84, for Spanish Final Interpreter Readiness test, I believe.

For the impact of training, we identified people within our rather small number of participants who had forty hours of training or less, and fifty hours or more. We wanted to know whether there was a difference in their performance on the test, because we thought the results would give us an idea of whether the tests were measuring the appropriate skills, knowledge and abilities. We found that on both the Initial and the Final Interpreter Readiness Tests there were no significant differences in performance. We also investigated whether amount of experience would have an impact on test performance. We thought people who had more experience would do better on the tests if the tests really measure valued and valuable skills. We identified people who had had one year or less of experience in our sample and people who had one year or more. We found on the Initial Interpreter Readiness Test that there was a statistically significant difference. Some of you might say, "Yeah!" Whoops! It was the less experienced people who did better! I have to say, I don't think this finding is a reflection on the test. My feeling is that the test is soundly designed, and we simply had huge problems in getting people to come and take the tests. I am quite certain that those people who did participate are not representative of the typical audience for the test. There was no significant difference for experience on the Final Interpreter Readiness Test.

Because of the very small sample for the different combinations of tests, and because there are questions that remain, we still need to collect validity evidence for the tests. There is a second phase of validation work that has been funded. I'm not going to go into the details of that; instead I would like to share some more thoughts on the validation work that has been done. First, the empirical basis for a test may provide convincing evidence of its content validity, but collecting evidence to support the usefulness of a test, that is, its value for screening purposes, is challenging, as is investigating the meaningfulness of test scores. What is a good score and at what score can one be assumed to be qualified or a master? We tried to investigate these questions through the differential group studies and found that for these people who were willing to come in and take the tests, there was no apparent difference. So when thinking about recruiting field test participants, we need to consider how a representative sample can be obtained. I ask myself, "Who likes to take tests? Who is willing to come in and do two hours of testing for sport? And are those people typical?" I'd like to suggest that they're not, which is a huge problem when we're investigating the validity of a test. We have to be very careful to design the research so that we do have access to people

who are typical of the intended audience. I don't think the people that we recruited were, even though they were offered fifty dollars and our admiration and esteem. Is fifty dollars enough to alleviate test trepidation for most people? I don't think so. There are many other theoretical and logistical and practical concerns related to designing and conducting validity research that one has to think about, ponder on, work collaboratively to solve before beginning the really important part of the development of the test.

Q: You mentioned the Cantonese test. There is a written part, right? Is the language itself Cantonese or Mandarin?

A: You know I believe it's Cantonese. At this level of questioning, we would have to look at the test itself. It was reviewed by a person as part of the validity work that I was involved in. We reviewed all of the tests and I was told that it was Cantonese. And I have been told by the people who developed it that it is Cantonese. The fact that it was reviewed by someone who wasn't involved in the development and who also assured me that it was Cantonese was comforting, because when we had the Hmong test similarly reviewed by someone who had not been one of the informants in its creation, we discovered that in fact it was not the variety of Hmong that we had been informed that it was, not precisely the color of Hmong that we had been informed.

Q: You mentioned that the folks who had less experience did better on this test. How was experience defined?

A: Experience was defined as they answered on a questionnaire saying, "How long have you been working as an interpreter?" And if they identified more than a year, we put them in one group.

Q: Is this regardless of actual experience within that year, because I've been doing it for a year, but I've only interpreted five times.

A: Just how long they've been doing it. The question was, "How long have you been doing interpreting?" People said, "I haven't done it." People said, "Two months." People said, "I've been an interpreter since I was twelve." So just simply on the basis of this self-report data, we tried to sort people out into people who had had more than a year and people who had less. And that could be one of the reasons that we came up with funky, unexpected outcomes.

Q: Along that line, did you ask any questions in regard to the fact that they had formal training or not? I'm just wondering whether that may have made a difference.

A: Yes. We did an analysis for training and people who had had forty hours of training or more were put in one group, and people who had forty hours or less were in another. Forty hours was defined because that is typically the length of the training program that is offered by the Connecting Worlds. By the way, during the eighteen months that we were working on this project, that training wasn't offered. So we couldn't actually do this project pre-post as we had intended to do. So we did have a group of forty hours or less and more than forty hours. More than forty hours might have been "I have a bachelor's degree. I did a certificate. I did this training, plus I've also had another six months of study somewhere." And I would say that the people who identified themselves as having

forty hours or less had done the training that was offered by whichever organization they were being tested at.

Q: Did they fare better or worse?

A: They were the same. No difference in the groups.

Q: I have one comment and then a question. In the early stages of the federal exam, back in 1980, we did a study of the first group that took the test and found a negative correlation between years of experience and passing the test, and no, there's nothing wrong with the test. What's wrong is that people were working who did not have linguistic and interpreting capability at the level required and probably didn't understand the goals of interpreting, fidelity, concept verbatim, etc, etc. They didn't understand their ethical roles, and that's certainly a validation of what you found.

A: It's possible. I have a feeling. I hate to even call them findings, because they're not really findings. They are based on a very, very small sample of people who were recruited. Thank you very much for recruiting them. You know, people worked very hard at these sites to get people in so that we could try these tests out, but those people are not typical of the general group of community medical interpreters that would come into those offices.

Q: Well, these were real findings, and we had like two-hundred-fifty people. And it was really shocking to everyone at the time, and we were all aghast, and we had to figure out what was going on.

A: I appreciate that, and as we do more research on this test, if we continue to find this profile, I think it will have really important implications for attention to assistance for community medical interpreters.

Q: Right. Exactly. And then, number two, my question is you found no increase, or significant increase in the testing after taking the Connecting Worlds?

A: None of the participants actually took the training. That was our original design, to do this as a pre-post, to have people take the appropriate language proficiency test and the English proficiency test, and the Initial Interpreters Readiness Test, and then engage in the training, and then do the post test. But that did not happen at any of the sites, because during the eighteen months we were engaged in the project, none of the sites offered that training because they didn't have the funding at that point in time to do it.

Q: But you reported it as a finding? Did I read it wrong? There was no significant increase.

A: That is for experience and for amount of training, but it wasn't the training that was provided by these agencies. It was the self-report training. Thank you for clarifying that.

Q: If that were true, then there would have to be a real study of the tasks, skills and abilities being taught in that program and if they matched the testing, which I wonder about.

A: In fact, when Claudia was engaged to develop the test, it was intended that it would be used with the Connecting Worlds training. But the Connecting Worlds training was not done when this test was done. So in fact, she had to kind of imagine and hypothesize, and I understand that it is now a curriculum which is used in different ways at the different organizations depending on the type of people that come into these five different organizations. I'm sorry about that lack of clarity. There was no training done with these participants as part of this study.

Q: On the Hmong interpreters, was your experience that you were just not able to get anybody to come and take the test? Was that what happened?

A: For the validation project, we didn't work with the Hmong test. We focused on the Spanish and Cantonese tests because in the initial development of the Hmong test, I think that there were some issues. I wasn't a complete insider on that first project, but I did work with the Hmong group in the orientation and training we did to introduce the tests to those people. And I believe that the nine interpreter trainers who were working for one of the organizations and had contributed to the development of that test also took the test. I think that there were a couple of other people who took the test as part of this preliminary pilot, and none of them completed the screening. And when I asked them about it, because I was the person who was working with that group, they just said, "Oh, we don't do that."

Q: My comment would be that in Madison, when we actually started to have an assessment for our Hmong interpreters, the same way we were doing for Spanish interpreters, all of them refused to come and take the test. As funny as it sounds, it really meant something. So I would like us as we think about this issue of certification to keep in mind that we're not going to be able to go and drag people from their homes to come and take the certification. So when I was talking about the end goal being language access for limited English-speaking people, we really need to keep those issues in mind. If all my Hmong interpreters in Madison refused to take the test, then I can have as many tests as I want, it's not going to happen. And they told me they just don't think the test is a good way to assess their skills and that they were insulted by even the idea, thus they were not going to come. Just food for thought.

Lessons learned from current state efforts

Summaries of Invited Reports on State Initiatives

Enrica Ardemagni, Indiana Commission on Health Care Interpreting

The Indiana Commission on Health Care Interpreters was formed because of legislation prompted by lobbying from the non-profit advocacy group Indiana Minority Health Coalition. While the first failed legislative directive in Indiana called for certification of health care interpreters and translators within a year's time, the current mandate forms the Commission's four goals: to define health care interpreters and translators, to publish standards of practice, to determine regulatory oversight, and to work on training and qualifications.

The fifteen-member Commission was required to include representatives from higher education, minority health, the state medical association, and county health departments. Other required representatives included an interpreter supervisor, two advocates, a representative from the Indiana Health Professions Bureau who could discuss certification versus licensure, a representative from the Department of Corrections, a member of a professional translation or interpreting organization, a member of the State Department of Health, a representative of Medicaid, a hospital representative, and a member of the Department of Education. The commission's bylaws also allowed for an unlimited number of ad-hoc appointments. Ad-hoc appointees brought representation from the deaf and hard of hearing community and greater representation from professional translators.

Forming sub-committees, the Commission researched professional definitions and qualifications of health care interpreters and translators. Fifteen members and invited outside experts committed to three months of intense work preparing a report to present to Indiana legislators and to the Indiana State Department of Health in October, 2004. The Commission relied on expertise from other states, encouraged the dissemination of processes for certification, advocated the adoption of standards of practice as well as a national code of ethics, and proposed a five-year implementation of certification on the state level. In taking these steps, the Commission remained hopeful that in the interim, a national certification would emerge to supercede state efforts.

The Commission did provide definitions for health care interpreter and health care translator. For the benefit of legislators, extensive terminology was included as background information. Members of the Commission have been promoting their work at town hall meetings, workshops, and conferences around the state. Like all state offices in Indiana, the Office of Minority Health is highly influenced by prevailing political winds. Nevertheless, the current State Health Commissioner endorses the Health Care Interpreting and Translation Commission's work and she advocates for the creation of a registry of health care interpreters. Meanwhile, having read the report compiled by the Commission on Health Care Interpreting, Indiana's governor now calls for the establishment of an educational component to further the certification effort in the state.

Maria Michalczek, Health Care Interpreting Legislation in the state of Oregon

Note: Dr. James Mason (Director, Office of Multicultural Health for the State of Oregon Department of Human Services) is co-author of the materials presented by Ms. Michalczek at the Expert Panel meeting.

Despite common beliefs, linguistic diversity does exist in the state of Oregon. In recognition of this fact, a 1997 certification law was proposed to bring certification to any profession that required interpreting (i.e. certified Spanish-speaking mechanics or certified Japanese-speaking home inspectors would exist). This law did not pass. A Racial and Ethnic Task Force uncovered that health care interpreter services were an important issue in the community. Committees met in 2000 and in 2001 to explore the procedures for qualifying and for certifying interpreters. These committees passed their findings on in the form of a bill to the governor. Oregon's governor is also a physician, and he clearly understood the need for certification and signed the bill into law.

Between 2000 and 2006, a twenty-five member council was assembled to oversee certification efforts. This council represented many different organizations, like Asian communities, women's groups, and others. Such a large council has proved difficult to manage. In 2007, at a Ways and Means Committee presentation, Michalczek explained how the group's efforts have been halted because of lack of funds. This council was comprised solely of volunteers who wrote comprehensive administrative rules outlining certification and qualification of health care interpreters. The Office of Multicultural Health is hopeful that the legislature will grant it the funding it needs to move forward. The formation of the governor-appointed Council represents the most significant component of Oregon's law on health care interpreting. The Council's volunteer efforts focus on definitions to determine interpreter qualifications and on oversight of the creation of an interpreter registry (for summer 2007). The council wishes to know who is out there and what they are doing as interpreters. The notion of a registry must be defined as well.

The Health Care Interpreter Law recognizes two classifications of interpreters, (qualified and certified) and both require 60 hours of training, state orientation, and language testing. Certification requires passing an exam. Administrative Rules for qualified and certified health care interpreters were approved in November, 2006. The qualification process and the registry are being rolled out in the summer 2007. The certification piece is lagging behind because of the complexities of devising a statewide examination. Among other lessons learned, Michalczek cites the need for careful and rapid review of the many drafts of proposed legislation, the need to appeal to varied, often monolingual supporters of interpreter legislation, and the need for community input at all phases of the process, especially in rural settings. Challenges include working "ahead of the curve." Many stakeholders do not yet understand the issues at hand, and this slows down the push for the passage and implementation of legislative reform. Budgets and funding are ongoing challenges in the state of Oregon.

Armando Villareal, Iowa Division of Latino Affairs

With \$250,000 federal funding secured, the Iowa Division of Latino Affairs was able to create a two-semester course at Des Moines Area Community College to qualify Spanish health care interpreters. A curriculum was already in place, and the Division of Latino Affairs moved within four months to write the Administrative Rules, therefore the program has graduated its first class. Students enrolled in the program receive college credits and are placed on a statewide roster that lists their training, qualifications and availability.

In Iowa, interpreters are not certified. If an interpreter meets minimal requirements, he or she is called “qualified,” both by the state and by the Division of Latino Affairs. The Des Moines Area Community College’s first cohort consisted of 42 students recruited, of whom 29 enrolled and 26 graduated.

Funding secured by the Division of Latino Affairs pays for program tuition, a bilingual dictionary, one staff member, and program publicity. The official website provides specific information on courses. Villareal congratulates his colleagues in Iowa for moving quickly and independently to create training for interpreters with the assistance of funding from the Iowa legislature. Rapid recruitment occurred through five Spanish language newspapers in Iowa, through radio broadcasts and through a state-sponsored list serve.

Carol Berg, Interpreting Stakeholder Group (ISG) in Minnesota

In October 2003 the Upper Midwest Translator and Interpreter’s Association invited stakeholders to meet to consider the spectrum of needs for improving interpreter services in Minnesota. Those at this first convening saw a need for continuing this effort, and out of this interest the Interpreting Stakeholder Group was formed in 2004, with the goal of improving delivery of interpreting services, promoting the “professionalization” of interpreting, and working toward equal access for LEP populations. The ISG has the following objectives:

- Identify and promote training opportunities
- Identify and implement internship opportunities to increase and enhance experience of interpreter trainees
- Improve evaluation of interpreter qualifications and performance, including standardized tests, consumer feedback and user feedback
- Explore options for enhanced interpreter service delivery to ensure access
- Seek funding opportunities to expand training, evaluation and partnering with other stakeholders
- Enhance the profession by advocating for appropriate remuneration and adequate employment opportunities
- Educate entities that rely on interpreter services on how best to work with interpreters
- Work with the legislature to improve the delivery of training and services for interpreters and their clients.

Members of the ISG include interpreters, trainers, academicians, interpreter services agencies, health care organizations, health and human services' representatives, local social services agencies' representatives, health education specialists from rural Minnesota, and a representative of the Minnesota State Court Interpreter Program.

The following are some of the ISG's major accomplishments. In monthly meetings, they share expertise, invite experts to give updates on education and service delivery and to describe what is taking place in Greater Minnesota. They have organized training in the Twin Cities metro area and in rural settings. They have devised a "Business Case Presentation" to attract funding agencies that they also invite to join the ISG. They have sponsored presentations on the current status of certification efforts and on interpreting in state courts. They sponsored a talk by Laurie Swabey on internships for sign language interpreters in order to investigate models for spoken language interpreters who also need internship opportunities. The group devised mentoring models to propose to employers with interpreters on staff, and finally the ISG secured grants from the Bush Foundation and the Bremer Foundation to conduct ethics and orientation training in Worthington MN, to offer an Introduction to Interpreting course via ITV, followed by a Consecutive Interpreting course. These grants covered enrollment costs for participants from Willmar and Crookston, Minnesota. Training in both the Twin Cities metro area and rural Minnesota is targeted to interpreters of Somali, Hmong, Oromo, and Spanish.

The ISG wants to convene a broad cross-section of partners to really improve delivery of interpreter training and services. While the current focus is on health care, there is a desire to move into the areas of mental health, law, education, child protection, and social services. The group has identified strategies for improvement and targeted growth, focusing on education, training, credentialing, and regulation of the field of interpreting. The group has made efforts to improve networking in order to support the role of the interpreter. The Internet Stakeholder's Group is also interested in formalizing its own status as a means of furthering its agenda.

The three priority areas that the ISG has identified include more educational opportunities to improve the quality of interpreting in Minnesota, increased funding to subsidize training costs, and the development of an interpreter roster. Having recently reached an agreement with the International Institute of Minnesota to serve as fiscal agent, the group will be able to write grants and manage resources more easily.

The group will continue to seek funding in order to provide many more training opportunities. On the certification front, the ISG seeks to share in the effort to develop a national certification process and to devise procedures for training certification raters. The ISG proposes to pool funds and to join forces to invigorate a newly formed Interpreter Services Work Group, mandated by the state to determine how to fund and distribute services in an appropriate manner. Its report will be considered by the state legislature in January 2008.

Questions Posed to Workgroups

Questions concerning interpreter competencies:

1. What skills and knowledge should a certified interpreter be able to demonstrate in an exam?
2. Should there be certification requirements other than passing an examination and, if so, what should they be?
3. What steps need to be taken, over the next six months, to move towards certification or towards answering the previous two questions?

Questions concerning test development and administration:

1. Who would have the capacity and ability to develop a certification test?
2. What research is needed before a test can be developed - when could it be developed?
3. What steps need to be taken, over the next six months, to move towards certification or towards answering the previous two questions?

Questions concerning implementation:

1. What are the steps to implementation (and the associated timeframe)?
2. Should a roster/registry be part of the preparations for certification and, if so, should standards for inclusion on a registry be created for individual states to use?
3. What steps need to be taken, over the next six months, to move towards certification or towards answering the previous two questions?

Workgroup Reports

I. The Relationship between Individual States and National Initiatives

Group members: Armando Villareal, Maria Michalczek, Mauro Yanez, Carol Berg, and Enrica Ardemagni (Berg and Ardemagni, facilitators)

Note: “State” doesn’t necessarily mean a public or government agency; it may be a state or regional organization.

How can states work with the national initiative?

1. States can use common definitions for words like “registry.”
2. States can share a model for “registries” (software) developed by CHIA
3. States can compile list/descriptions of currently available training, and make recommendations re what should comprise training. (NCIHC is planning on working on this, contingent on funding.)
4. States can serve as pilot sites for training, certification tests, etc.
5. National body could centralize/coordinate state initiatives.
6. National organization can compile a national registry of certified interpreters in each state.
7. National organization could provide model statutes or legislation (RID or NAD has model language—check this).
8. National group can recommend a model training program, drawn from actual existing state programs
9. National group can help states, not duplicate efforts, e.g. to recruit trainers, have a rotating train-the-trainer institute
10. National body can offer language to help states build a “business case” for interpreters to show how trained interpreters are cost-effective. NOTE: NCIHC annotated bibliography of research to be placed on website www.ncihc.org could be source of an issue paper.

How can states work together?

1. States may organize individual state databases or registries of interpreters. A “registry” may vary in complexity from state to state. In some cases it may just be a database of unverified information. For us, the word “registry” means a listing of people who meet some minimal qualifications and are therefore authorized by the state to interpret or to have their services reimbursed.
2. States can develop training programs that can inform national training and cooperate with one another.
3. Individual states can get started in preparing interpreters to eventually become certified.
4. States should continue to meet face to face (as in this forum).
5. States could join together to sponsor train-the-trainer institutes.
6. States can develop tests in various areas or for different languages.
7. States can collaborate on research to build a business case – or compile such research that exists.

II. Interpreter Competencies and Test Development & Administration

Group members, Interpreter Competencies: Karin Ruschke, Isabel Arocha, Elizabeth Nguyen, Roseann González, Laurie Swabey (Bruce Downing facilitator)

Group members, Test Development and Administration: Hungling Fu, Jean Turner, Frances Butler, Nataly Kelly (facilitator)

The two work groups on Interpreter Competencies and Test Development had determined that it would be beneficial to work together, so that members of each group could learn from each other. They began their joint sessions by developing a working statement to describe the overall purpose of the certification program. The group agreed that the statement would evolve over time, as the details of the program are concretized. The draft statement is provided below.¹

Draft Working Statement

The purpose of the health care interpreter certification program is to certify that a person has demonstrated the essential competencies* to interpret in health care settings.

* It was noted that the term “competencies” would need to be defined in greater detail at a later stage, to include the specific areas of competency to be evaluated.

Using this working statement, the group discussed potential components for the certification program.

The group agreed that it would be important to have a “battery” of tests. In other words, it would not be a single test, but rather, a series.

Component A

Professional Standards and Ethics Test

This test would evaluate knowledge of ethics and standards of practice.

This test would present scenarios to the candidate verbally.

Candidates would also provide their responses verbally (ideally). The team would like to avoid written tests, since these may present challenges for some groups, and because written skills are not of primary importance to interpreters.

The test could be delivered entirely in a web-based or virtual environment, with an audio and/or video component.

The test could allow for participants to respond in either language.

It was noted that the verbal responses would make scoring the test quite expensive.

This test would last for approximately 15 minutes.

Professional Standards and Ethics Training

The importance of preparing candidates for success on the test was also discussed. For this reason, the possibility of providing an (obligatory) training (prior to the test) was discussed. Ideally, the training on standards and ethics could be delivered online, in a web-based environment. The training could then be delivered free of charge, or at a low

¹ We are indebted to Janet Erickson-Johnson for moderating this working session and to Nataly Kelly for submitting her own notes from the session in a form that provides the basis for this section of the report.

cost, to motivate candidates to complete it. (It was noted that sometimes when no training exists for a given subject there is a “backwash” effect – trainings may start to pop up in the field to cover the content area to be tested; however, a negative side effect can be a lack of uniform standards for training and/or quality control.) To provide some degree of uniformity and ensure that candidates have an opportunity to learn the information on which they will be tested, some type of standard training module would be highly desirable. The group was undecided about whether such a training should be mandatory or optional. One possibility would be to provide candidates with a “Checklist” of recommended steps for certification. Prior to taking the Professional Standards and Ethics Test, it would be recommended that candidates for certification complete the virtual training module.

- It was also noted by the group that certain standards of practice can be successfully evaluated in a performance test, as well as in a knowledge-based test. For example, a test could be designed to include certain utterances that are much longer than the typical utterance that an interpreter would be required to interpret. Even in a web-based test, the interpreter would have the ability to request a repetition and have the utterance repeated.
- It was also pointed out that, in natural speech, humans do not usually repeat things exactly as they said them the first time. If a repetition is requested, the utterances could be recorded several ways, to ensure that this is addressed.
- Ideally, the Standards and Ethics test would cover the bulk of content in the areas of Standards and Ethics. The performance test, on the other hand, would likely include a few specific key standards that could be tested in a role-play format (e.g. requesting clarification/repetition, pre-session). Others, such as controlling the flow of speech, would be difficult to “test” in an automated or pre-recorded role-play format.

IMPORTANT: For all of the skills tests listed below, the content would be driven by a prior Job Analysis.

Component B

Skills Test I – Role-Play Scenario

This skills test would entail a typical role-play scenario, and would be the section of the test that most closely approximates “real-life” interpreting.

The content area would most likely be primary care, which is common enough to all language groups, and which also would contain terminology that would be highly reflective of real-life interpreting work. Patient utterances could include typical slang and colloquial speech. Provider utterances could include technical medical terms, as well as jargon and provider colloquialisms.

This test would include scoring units that would enable examiners to measure terminology knowledge, grammar, accuracy and completeness, memory retention, etc.

This test could be delivered live, for greater realism, but would probably be delivered using recorded (audio or video) roles of provider and patient, to save cost and to insure greater uniformity for scoring.

This test would have a duration of approximately 12-15 minutes.

Component C
Skills Test II – Sentence Conversion

This skills test would entail conversion of sentences, in both directions (e.g., English into Vietnamese and Vietnamese into English).

This would be the area in which the largest amount of terminology knowledge can be assessed. It could cover a wider range of specialties than possible in the role play.

The sentences used would still be ordered in a way that gives the candidate some sense of logic and ability to understand the context.

This is the portion of the test that may include multiple “technical terms” in a single sentence, to truly assess terminology knowledge.

“Technical terms” includes both provider speech and patient speech. In other words, slang words for body parts, culture-specific remedies, names of medical procedures, medications, and other such terms would all be potential scoring units.

This test would have a duration of approximately 10-12 minutes.

Component D
Skills Test III – Sight Translation

This skills test would entail conversion of sentences from written form to verbalization in the other language.

It was pointed out that legal documents, such as consent forms, would not be good documents to use, because the language is easy to memorize, and therefore does not really test interpreting skills. In addition, the standards of practice dictate how such documents should be handled, so this could be covered under Component A.

The team noted that the types of documents would be determined by the Job Analysis. It is likely that many working interpreters are actively interpreting consent forms, even though the more cost-effective and ethical solution is to ensure that these documents are translated in advance rather than sight-translated. However, even if a Job Analysis does show that interpreters are interpreting complex legal documents such as consent forms, the test will be designed not to include such texts, since interpreters can decline to sight translate such documents, in keeping with the code of ethics and standards of practice.

By way of analogy, a “job analysis” of driving may show that the majority of active drivers fail to use a turn signal. That does not mean the driving test does not stipulate that drivers should use a turn signal.

In summary, it will be important for there to be a balance between the ethics and standards for this emerging field and the Job Analysis that shows what interpreters are actually doing, to promote higher standards of quality and better service for LEP patients.

The foregoing describes just one example of what the team predicts is likely to be found through the Job Analysis – that working interpreters are currently not following uniform ethics and/or standards. The certification process may serve partially to help establish some uniformity in this regard, especially if a training module is included.

This test would have a duration of approximately 6-8 minutes.

Component E
Skills Test VI – Simultaneous Interpreting

This skills test would measure simultaneous interpreting skills.

This test would be an **OPTIONAL** component, not required for certification, but optional for interpreters who wish to have a simultaneous “endorsement” listed in their credentials as a medical interpreter.

While not commonly addressed by training programs for health care interpreters, most working interpreters find that they are required to perform simultaneous interpreting, often in emergency settings, mental health settings, and also commonly, when interpreting for health education scenarios.

Having this component as an optional module would support two purposes: (1) it would enable interpreters with this additional skill to showcase their abilities and have formal recognition of this, and (2) it would begin to introduce the importance of simultaneous interpreting skills to the health care interpreting field at large, so that interpreters would begin to see simultaneous interpreting as a common part of the job. There is some intimidation regarding simultaneous interpreting, even though many interpreters, when trained, often find that it makes their job easier.

As this test is not a required component, approximate duration was not discussed and would need to be determined through further discussion.

Given that part of the goal of an optional simultaneous interpreting test would be to introduce this to the field, a desirable component would be a virtual training module in simultaneous interpreting. Similar to the Professional Standards and Ethics Training, this could be delivered either for free or at a low cost via the Internet.

Component F
Language Proficiency Test

There was unanimous agreement that language proficiency testing is important. However, there was discussion about when testing should occur. Some felt that language proficiency testing should be a prerequisite to training, as this saves candidates from wasting money taking a training that will ultimately be of minimal benefit to them.

Others felt that language proficiency testing should be a prerequisite to other performance testing, such as Components B, C, D, and E.

The timing of when language proficiency testing should occur, and whether or not it should be part of a certification program, or simply a prerequisite requirement, are issues which require further discussion.

It was also noted that several language proficiency tests already exist in the marketplace in a large number of languages. But they would need to be reviewed to determine which are appropriate for testing the proficiency of (prospective) interpreters.

Important Next Steps

The Test Development / Competencies Teams identified the following key tasks / action items that would be important to undertake in order to prepare for test development:

- Conduct a detailed Job Analysis of health care or “medical” interpreting. First, a decision will need to be made as to how to define the field: does it include home health visits, dentistry, mental health assessments, etc.
- Conduct a review of the National Standards of Practice and other standards publications, to determine which items are reflected in the Job Analysis, and which items from the Job Analysis reflect a lack of concordance with the standards, so that these items might be addressed through training and testing components (e.g., sight translation of legal documents, as described above)
- Conduct surveys to determine common content domains for various language groups, and those that are common to all interpreters in order to select content domains for role plays.
- Conduct a legislative review to ensure that the certification process reflects applicable law.
- Hold focus groups/expert panels to determine the specific competencies, i.e., the knowledge, skills, abilities and tasks (KSAT), to be performed by professional medical interpreters, and therefore evaluated. Conduct a review of literature and gather empirical data, such as job descriptions, information on error analyses, etc.
- Draft a list of desired test preparation materials (or develop materials that would be needed).
- Draft a checklist of “steps toward certification” for prospective candidates.

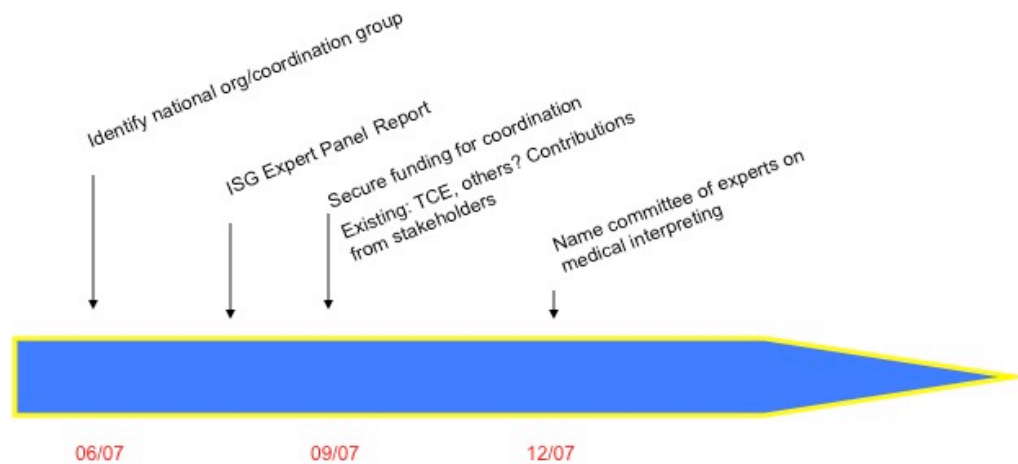
III. Implementation of a Process: Next Steps²

Group members: Maria-Paz Avery, William Hewitt, Shiva Bidar-Sielaff, Janet Erickson-Johnson, Cindy Roat (Patricia Ohmans, facilitator)

The Implementation Committee proposes a series of implementation steps, the first of which are specified on the timeline on **Slide 1 (next page)** that encompasses only the first six months. Step One in the process is to identify some national organizing or coordinating group. If we are serious about actually getting things done, there has to be some body to take the lead. For example, if you were going to hire a staff person to see that things get done, that person is going to have to work for some entity, what we have identified as this national coordinating group. The next step is to elicit a report from this Expert Panel — presumably conclusions, feedback, and recommendations. Then obviously nothing happens unless you secure funding, and so by September 2007 we’re suggesting that funding may come forward There’s a reason for picking that date, as explained below. Then the organizing/ coordinating group will need to identify a panel of experts in medical interpreting that would then begin to do or to oversee the sorts of things that the Competencies and Testing workgroup have identified.

² This section of the report is based on the audio recording of the workgroup report on Implementation given orally by William Hewitt and Shiva Bidar-Sielaff accompanied by PowerPoint slides prepared by Shiva Bidar-Sielaff for the workgroup.

The next six months



For the period after six months (**Slide 2, next page**) the Implementation workgroup offers a bulleted list of other things that would need to be accomplished, such as compiling existing information on certification issues, what job analyses already exist, and what we know about adequate proficiency levels. There will need to be a process to select the actual certifying body. Suppose we already have a test that can be used; how then does certification take place--what would be the certifying entity. But actually before there is an actual test, there will need to be a template for test design, and there will need to be an opportunity to get feedback on that, and then on a draft test for a particular language pair, and so on—a continuous iterative process. So a process needs to be put in place for the steps of design, testing, and review. Another task will be exploring ways to market the certification process to the stakeholders: getting the medical establishment to buy into the value of certification, and to understand the lengthy iterative process. Along the way it will be necessary to secure funds for the actual test development. To make the process collaborative and to obtain buy-in, there's a need for additional steps, such as conducting forums, obtaining and analyzing survey feedback, again furthering the iterative process, The workgroup stopped at this point (this is after all just a first step in designing and elaborating a possible process), but obviously there's much more to be done.

After six months

- Compile existing information on certification issues, e.g. job analyses, adequate language proficiency levels
- Select certifying body
- List was to get feedback on test as it is
- Explore ways to “market” certification process to stakeholders
- Report to stakeholders-iterative process
- Secure funds for test development
- Conduct forums, survey, feedback –iterative process
- Develop pilot test

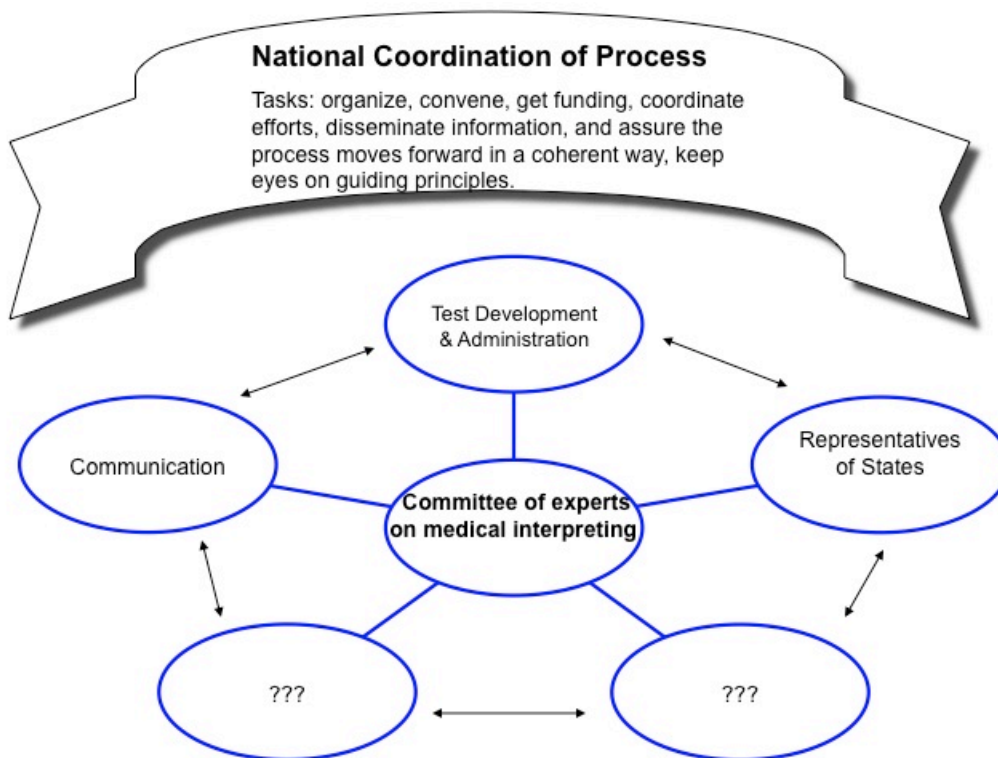
In the meantime, while this process is hopefully moving forward, other related activities will also be underway, **as shown on the Slide 3**. The development of a set of national standards for health care interpreter training and education is being planned for the near future, and it seems likely that at the state and local level both governmental and non-governmental organizations will be developing some of the products that we have talked about here: rosters and registries, guidelines for recruiting, language screening programs, interpreter training/education programs, and perhaps a template for language access legislation—a model statute that could be proposed in state legislatures.

In the meantime

- **National**
 - Standards for Health Care Interpreter Training/Education
- **State/Local**
 - Rosters/registries
 - Guidelines for recruiting
 - Language screening programs
 - Interpreter training/education programs
 - Template for language access bills

The graphic on **Slide 4** depicts a model for the national coordination of the process. Recall that our recommended first step was to identify a coordinating group. The responsibilities of this group include organizing, convening, getting money, coordinating efforts, disseminating information, and assuring that the process moves forward in a more coherent way than any of the scattered past efforts. Equally important for our model is what this group is NOT doing. In this model the coordinating organization would NOT be making judgments about competencies to be tested, or test design, or the eventual administration of certification testing as they are organizing the process. Their responsibility would be coordinating what needs to be done for test development; they would be bringing representatives of stakeholder organizations on board for collaboration and communication; they would be convening the committee of experts on medical interpreting and competency testing. It is THAT group--the committee of experts--that would need to make some decisions. And the representatives of state organizations and other stakeholders—people who are buying into the process—are going to be participating in a decision-making process. And there are other tasks that will be the responsibility of the committee of experts, some of which we haven't tried to specify yet.

Slide 4



There will need to be subcommittees, such as the group of representatives of states (and other stakeholders), and the group specifically responsible for test development. “Communication” on the chart indicates the need for regular communication among all the parties involved. Communication might be the responsibility of a subcommittee. As a simple example, a listserv might be an efficient way to promote communication, to keep stakeholders and organizational representatives

involved, to get good input, but somebody has to create the listserv, manage it, and so on. So that's part of what we mean by communication.

Now, going back to the beginning of the process of implementation, the very first essential step is the identification of who will get the process started. The Implementation workgroup was asked to suggest options, wherever we could, so the group looked at options for the national coordination organization, aware that a coordinated process is impossible until some agreement is reached on this first step. It's far from sufficient, but it's necessary. So the Implementation workgroup considered three options, as shown on **Slide 5**.

Slide 5

Options for national coordination organizations

Option 1: Form a new coalition with representation from CHIA, IMIA and NCIHC

Option 2: Ask states to form own stewardship coalition

Option 3: NCIHC

Option 1 is to form a new coalition or taskforce with representation from major interpreter and stakeholder organizations such as IMIA, CHIA, UMTIA, ATA, and others. It was not clear to the workgroup how this could be done--who's going to form the coalition—but that's an option. Option 2 is to ask the states or state organizations to form their own stewardship/coordination coalition. Again, it is not clear who would do the asking or how this would come about: someone would have to bring the various state organizations together and make this happen. Option 3 is for an existing organization, the National Council on Interpreting in Health Care (NCIHC), to do the initial organizing and fill the coordinating function—to convene an expert panel or taskforce, with stakeholder representation that would make the decisions.

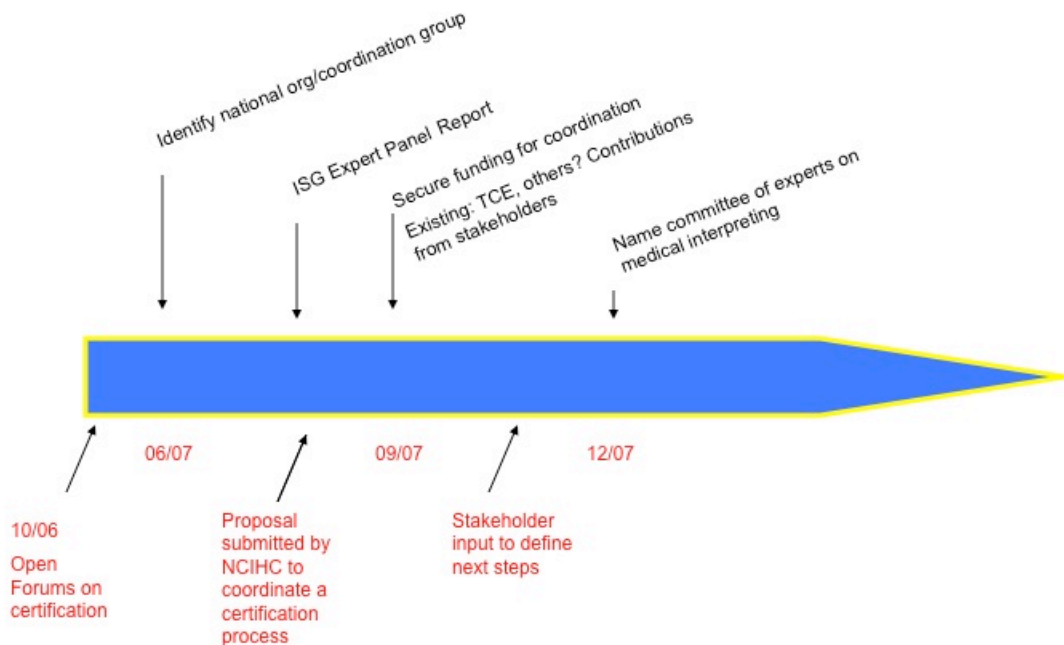
The organization chart presented earlier shows how responsibilities would be divided between the coordinating group (“national coordination of process”) and the expert group (“committee of experts on medical interpreting”): the first group being responsible for organization and administration of the effort and the second group for communication, decision-making, test development and a certification process. It seems essential from the outset to have someone who is going to get paid to coordinate things. Looking at the three options, and the fact that the first two options require new

infrastructures or organizations (to receive funding, to hire a coordinator, and to begin planning and bringing people together), the advantages of starting with an existing organization seem obvious. But this of course begs the question: which organization?

Looking at the timeline and immediate questions like who is going to coordinate the next meeting, the workgroup decided to place on the table a concrete proposal that, if adopted, would make it possible to move forward immediately. If agreement cannot be reached on a coherent agenda and who can be relied on to take the next step, then there is the prospect of uncoordinated and competing actions which presumably no one wants to see. So the workgroup wanted to lay out what the NCIHC has already done and is doing, and why it is in a position to move the agenda forward immediately. Then the workgroup would like to open the options up for discussion by everyone at the table.

To the time line **on Slide 1** we have added the steps that the NCIHC has taken thus far and has proposed to advance the certification agenda (**see Slide 6, next page**). First, there is the series of forums on certification at national and regional conferences that the NCIHC has organized beginning in October 2006, at the Quality Health Care conference in Seattle, at the MMIA conference in Boston, at the CHIA conference in San Jose, with others scheduled for coming months. Meanwhile, there have been other important meetings such as the May 1st forum in Boston, this one—the Expert Panel in Plymouth, Minnesota. From these forums and other sources there is already a body of knowledge that needs to be pulled together in a coherent way. The NCIHC has been asked by the California Endowment to submit a proposal for the funding of the planning and coordination phase of a certification process. This proposal will be submitted before the end of June 2007. Funding is not guaranteed, but it is significant that the foundation requested the proposal. So next steps have already been planned, as part of this proposal, and a decision on funding is expected as soon as September 2007. If funding is obtained, the first step would be for the NCIHC to convene a group of stakeholders and decide in a stakeholder meeting who should be on that committee of experts, what subcommittees or related workgroups would be needed, and what tasks would need to be assigned to different groups. In our discussions, the Implementation workgroup agreed that Option 3 offers the quickest way to move this conversation and process forward, because the NCIHC already has the capacity and recognition as a leader in his area at the national level, and already has things in place and moving forward.

The next six months



Note: For a summary of the Expert Panel’s closing discussion, centered on this Implementation Workgroup report, please see the “Expert Panel Overview” at the beginning of this document (page 9).

The slides that accompanied other Expert Panel presentations are available in PDF format on request. For this and any other inquiries regarding the Expert Panel or this final report, please email Bruce Downing: bdowning@umn.edu.

Additional copies of this report can be downloaded from the web site of the Upper Midwest Translators and Interpreters Association (UMTIA): <http://www.umtia.org>.

Appendix A

Agenda – Expert Panel on Testing and Certification of Community Interpreters

Day 1 (06/13/07)

7:30 – 8:00	Breakfast	
8:00 – 10:30	Plenary session	<i>Welcome, Overview of Expert Panel objectives, Lessons learned from national certification initiatives</i>
10:30 – 10:50	Break	
10:50 – 12:20	Plenary session	<i>Lessons learned from state-level initiatives</i>
12:30 – 1:30	Lunch	
1:30 – 3:30	Plenary session	<i>Lessons learned from current state efforts</i>
3:30-3:45	Break	
3:45 – 4:30	Workgroup session	<i>Implications of the first day's discussion for workgroup tasks</i>

Day 2 (06/14/07)

7:30 – 8:00	Breakfast	
8:00 – 9:30	Plenary session	<i>Corporate & non-profit certification exams</i>
9:30 - 10:00	Break	
10:00 - 12:30	Plenary session	<i>Full-panel discussion of questions concerning interpreter competencies, test design and administration and certification implementation</i>
12:30 – 1:30	Lunch	
1:30 – 3:00	Plenary session	<i>Pointing the way forward for Minnesota and other states interested in certification</i>
3:00 – 3:20	Break	
3:20 – 4:30	Workgroup & Plenary session	<i>Approval of workgroup agendas for Day 3</i>

Day 3 (06/15/07)

7:30 – 8:00	Breakfast	
8:00 – 9:20	Workgroup session 1	
9:20 – 9:40	Break	
9:40 – 11:00	Workgroup session 2	
11:00 – 11:20	Break	
11:20 - 12:40	Workgroup session 3	
12:40 – 1:40	Lunch	
1:40 – 3:30	Plenary session	<i>Final discussion and debriefing</i>

Appendix B

READING LIST ON CERTIFICATION

Primary Readings

This select group of readings is highly recommended.

Durley, Cynthia C. *The NOCA Guide to Understanding Credentialing Concepts*. National Organization for Competency Assurance, 2005.
<http://www.noca.org/members/CredentialingConcepts.pdf>

Kelly, Nataly. *Interpreter Certification Programs in the U.S. Where Are We Headed?* (*ATA Chronicle*, January 2007).
http://www.atanet.org/chronicle/feature_article_january2007.php

Mikkelson, Holly. *The Professionalization of Community Interpreting*.
<http://www.acebo.com/papers/profslzn.htm>

Roat, Cynthia. *Certification of Health Care Interpreters in the United States A Primer, a Status Report and Considerations for National Certification*. The California Endowment, 2006.
<http://www.calendow.org/reference/publications/pdf/cultural/Certification%20of%20Health%20Care%20Interpreters%20in%20US.pdf>

Secondary Readings

These readings, as listed in various categories, are also recommended.

Association Viewpoints

Hamm, Michael. *An Executive Summary: Review of the ATA Certification Program*. www.atanet.org/bin/view.pl/24113.htm.

National Association of Judiciary Interpreters and Translators. *Frequently Asked Questions about the National Judiciary Interpreter and Translator Certification (NJITC)*.
<http://www.najit.org/examfaqs.html>

National Council on Interpreting in Health Care. *Final Report on the Pilot of a Certification Project for Spanish-English Interpreters in Health Care (2003)*.

http://www.ncihc.org/NCIHC_PDF/NCIHC%20Certification%20Pilot%20Final%20Report.pdf

National Council on Interpreting in Health Care. “NCIHC and National Medical Interpreter Certification” (letter, May 1, 2007).

Academic Viewpoints

Chiaro, Delia and Giuseppe Nocella. Interpreters’ Perception of Linguistic and Non-Linguistic Factors Affecting Quality. *META*, XLIX, 2, 2004.

<http://www.erudit.org/revue/meta/2004/v49/n2/009351ar.pdf>

Clifford, Andrew. Discourse Theory and Performance-Based Assessment: Two Tools for Professional Interpreting. *META*, XLVI, 2, 2001.

<http://www.erudit.org/revue/meta/2001/v46/n2/002345ar.pdf>

Kurz, Ingrid. Conference Interpreting: Quality in the Ears of the User. *META*, XLVI, 2, 2001.

<http://www.erudit.org/revue/meta/2001/v46/n2/003364ar.pdf>

Moser-Mercer, Barbara, Ulrich Frauenfelder, Beatriz Casado, and Alexander Künzli. Searching to Define Expertise in Interpreting. *Language Processing and Simultaneous Interpreting*. Edited by Birgitta Englund Dimitrova and Kenneth Hyltenstam (John Benjamins, 2000).

Pöchhacker, Franz. Quality Assessment in Conference and Community Interpreting. *META*, XLVI, 2, 2001.

<http://www.erudit.org/revue/meta/2001/v46/n2/003847ar.pdf>

Court Interpreting Viewpoints

Hewitt, William. Court Interpreting Proficiency Tests: A Summary of What They Look Like and How They are Developed.

http://www.ncsconline.org/wc/publications/Res_CtInte_StateCrtJV20N1TestAtAGlancePub.pdf

National Center for State Courts. Consortium for State Court Interpreter Certification: Frequently Asked Questions.

www.ncsconline.org/D_Research/CourtInterp/Res_CtInte_ConsortCertFAQ.pdf.

Romberger, Wanda. *Skills Training for Foreign Language Court Interpreters: Does it increase the number of qualified interpreters?* National Center for State Courts, 2007.

Language Company Viewpoints

Geertsen Danyune, Nataly Romero, and Holly Mikkelson. *Testing for Certification: A Holistic Approach to an Interpreter Certification Program At Language Line Services*. Language Line Services, 2000.

Sawyer, David, Frances Butler, Jean Turner, and Irene A. Nikolayeva Stone. Quality Assurance Model for Remote Language Mediation (*ATA Chronicle*, August 2002).

ASL Viewpoints

National Distance Learning Center for Interpreter Education. Entry-to-Practice Competencies for ASL/English Interpreters.
http://www.unco.edu/doit/Competencies_brochure_handout.pdf

Registry of Interpreters for the Deaf. Overview of the Generalist Certificate.
http://www.rid.org/education/edu_certification/index.cfm/AID/45

Registry of Interpreters for the Deaf. Overview of the Specialist Certificates.
http://www.rid.org/education/edu_certification/index.cfm/AID/46

Registry of Interpreters for the Deaf. Overview of Testing Process.
<http://www.rid.org/education/testing/index.cfm/AID/83>

International Viewpoints

National Accreditation Authority for Translators and Interpreters.
Standards for Accreditation.
<http://www.naati.com.au/at-accreditation.html>

National Accreditation Authority for Translators and Interpreters.
Structure of NAATI Tests.
<http://www.naati.com.au/at-testing-procedure.html>

Stejskal, Jiri. *Survey of the FIT Committee for Information on the Status of the Translation & Interpretation Profession*. International Federation of Translators/Fédération internationale des traducteurs, July 2005.

Appendix C

PARTICIPANT BIOGRAPHIES

EXPERT PANELISTS

Izabel S. Arocha

Izabel is the President of the International Medical Interpreters Association. She is also the Cultural and Linguistic Educator at Cambridge Health Alliance, comprised of 3 community hospitals and 25 health centers. She oversees initiatives to increase the linguistic and cultural capacity and competency of the organization. She is an experienced consultant, trainer and group facilitator with a long history of multicultural advocacy. Izabel worked as a medical, conference, court and federal interpreter for many years prior to establishing her own translation firm, Global Mind, Inc. Her firm worked in many multilingual projects; for example, to develop the 30-language medical interpreter access poster required in all emergency rooms in Massachusetts. As an international presenter on cultural competency and medical interpreting roles and competencies, she draws on her own upbringing (having been raised in several countries). In addition to training and consulting, her skills include advocacy, curriculum development, program planning and management, marketing and media initiatives, community outreach and intercultural mediation. She is currently a lecturer of Boston University's one-year Medical and Legal Interpretation program and has developed and is on the faculty of the Mental Health Interpreting Certificate Program at Cambridge College. She holds a Bachelor in Management from Lesley University, a Translation Certificate from University of Cambridge, England, and a Masters in Education from Boston University. She is fluent in Spanish and Portuguese and speaks French well.

Maria-Paz Avery

Maria-Paz is a Senior Research Associate at Education Development Center Inc. in Boston MA where she directed a three year project to develop a 27-credit college level certificate program to prepare bilingual adults as medical interpreters. With members of the Massachusetts Medical Interpreters Association (MMIA), she helped develop its standards of practice which have now been recognized by the newly formed group, the National Council on Interpretation in Health Care, as the best available in the nation. She has been the chair of the MMIA Advisory Board, and worked with its certification committee in the development of a certification process for Massachusetts. She has presented at the International Conference on Community Interpreting on cultural issues in health care interpreting and the development of standards. Dr. Avery brings to this work over 25 years of training and consulting experience in the area of cross-cultural communication and the management of diversity in health, mental health, and educational settings.

Shiva Bidar-Sielaff

Shiva is the Manager of Minority Community Relations & Interpreter Services at University of Wisconsin Hospital & Clinics. She has worked extensively on issues of equal access to health care for limited English proficient (LEP) individuals on a national and local level. As the Co-chair of the Standards, Training and Certification Committee of the National Council on Interpreting in Health Care, she has been involved in developing the National Code of Ethics and Standards of Practice for Interpreters in Health Care.

Ms. Bidar-Sielaff is the Vice-Chair of the Dane County Latino Health Council and an active member of the Latino Support Network. In April 2000, Ms. Bidar-Sielaff was awarded the Dane County Public Health Leadership Award for Multicultural Health Care. She is the 2005 recipient of the Madison YWCA Woman of Distinction Award for her work in fighting inequality and eliminating racism.

Frances A. Butler

Frances A. Butler, Ph.D., is an Independent Educational Consultant in Language Assessment. Previously for 15 years, Dr. Butler was a Senior Research Associate and Language Testing Specialist at the Center for the Study of Evaluation (CSE) and the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) in the Graduate School of Education and Information Studies at UCLA. She has continued her association with CSE/CRESST as an independent contractor. Her most recent research efforts have focused on the articulation of academic English for broad application—curriculum, assessment, and professional development—in school settings. Known for her research in language testing, she has consulted on and directed language testing projects for such diverse organizations as the Los Angeles Unified School District, the Educational Testing Service (ETS), the Egyptian Ministry of Education, and the Kayenta Unified School District on the Navajo Reservation in Arizona. In addition, Dr. Butler has worked with *NetworkOmni Multilingual Communications*, a language services company based in Westlake Village, CA, on the development of their internal certification program for telephone interpreters. For six years, from 2002-2008, she served on the Technical Advisory Group for the *California English Language Development Test*, CA Department of Education. Currently she serves on an advisory board for the Test of English as a Foreign Language, ETS. Dr. Butler has presented on language testing topics at national and international conferences and has published in such journals as *Language Testing*, *Applied Linguistics*, and *Theory into Practice*.

Janet Erickson-Johnson

Janet is the Certification Manager for Language Line Services. She has overseen the development and administration of certification testing in the Medical, Court, and Insurance industries since January 2000 and has also played a lead role in the

development of basic interpreter training and Medical Interpreter Training for LLS, which she has delivered both in the United States and abroad. She received a Masters' Degree in Translation and Interpretation (Spanish) from the Monterey Institute of International Studies in 1994, and translated a book on ADD/HD from Spanish into English. Janet also completed a Medical Interpreting Internship at Stanford University Hospital in Stanford, California, in 1991 and subsequent to receiving her degree, taught Medical Interpreting at the Monterey Institute of International Studies. As a California Administrative Hearing Certified Interpreter (1990), Janet worked as a professional free-lance interpreter for 10 years, interpreting for medical appointments, medical-legal proceedings, court proceedings, administrative hearings, and in a variety of other settings. She has also been a presenter at the 2000 Massachusetts Medical Interpreters Association Annual Conference in Boston, MA, and the 2002 ATA Medical Translating and Interpreting Regional Seminar in Chicago, Ill., and is a member of the California Health Interpreters Association.

Hungling Fu (not available)

Roseann González

Dr. Roseann Dueñas González has been the Director of the National Center for Interpretation Testing, Research and Policy at the University of Arizona since its inception. She is responsible for overseeing the Center's operation, including the research, instructional programs that are developed and taught, financing and administering of the Center's several projects, and liaison functions between The University of Arizona, the AOUSC and other judicial bodies as well as national, state, and local government agencies and officers. Most importantly, she is responsible for the development of valid and reliable testing instruments, the training of oral examination raters, the administration of court interpreter certification exams throughout the United States and Puerto Rico, and the evaluation of these exams. In addition to these duties, she continues to be very active in research both individually and in collaboration with other scholars in the field. She also serves as a mentor to young scholars wishing to gain from her expertise. She represents the Center and the University at conferences and professional meetings in the field. Dr. González is a Professor of English and has been a member of the University Faculty for 25 years.

William (Bill) Hewitt

A veteran of work in the field of court administration, with experience throughout the United States and abroad. He began his work with courts as a juvenile probation officer in a small rural court in Washington state in 1973; he retired as Principal Court Research Consultant from the National Center for State Courts in 2007, with publications as author or co-author on a wide range of subjects, including delay reduction and case flow management, trial court performance standards, public defense, judge and staff workload

assessment, court reporting. Since 1990 his work has focused on language interpreting in the courts. His 1995 publication *Court Interpretation: Model Guides for Policy and Practice in the State Courts* is widely used and excerpted in court policy and planning documents, educational programs and court rules across the nation. Mr. Hewitt conceptualized and facilitated the establishment the Consortium for State Court Interpreter Certification in 1995. The Consortium is a voluntary collaborative effort among state courts that now includes 36 states and was a 2003 finalist in the Innovations in American Government Award Program sponsored by the Harvard University Kennedy School of Government.

He was the Project Director for the Federal Court Interpreter Certification Project from February, 2000 until his retirement. Mr. Hewitt received his MA degree in philosophy from the University of California (Davis). He was a Woodrow Wilson Fellow and he is a Fellow of the Institute for Court Management.

Nataly Kelly

Nataly is an independent consultant and researcher based in Nashua, NH. A certified court interpreter (Spanish), she is a former Fulbright scholar and author of *Telephone Interpreting: Telephone Interpreting: A Comprehensive Guide to the Profession* (Multilingual Matters, 2007).

Elizabeth Nguyen

Elizabeth is a native of South Vietnam. She has a Bachelor's Degree in French Literature and Philosophy, and a Master's Degree in Applied Linguistics. She is the Senior Diversity Services Specialist/ Operations Manager of the Language and Cultural Services Department at Children's Hospital Los Angeles (CHLA). Prior to joining CHLA, Elizabeth worked as the Cultural and Linguistic Training Specialist at L.A. Care Health Plan, the largest provider of Medicaid Managed Care Services in the country. Prior to that, she was the Program Manager at PALS for Health, a community-based organization serving the language needs of Asian and Pacific Islanders. Elizabeth's experience as a community and medical interpreter dates back to the early nineties. Her passion for advancing the health care interpreting profession and her commitment to enable equal access to health care for limited English proficient patients have led her through various arenas of activities that include program development, interpreter and provider training, consumer education, community outreach and advocacy. She is a Board Director of the California Healthcare Interpreting Association (CHIA), the former co-chair of the CHIA Standards and Certification Committee, and co-author of the book "*California Standards for Healthcare Interpreters – Ethical Principles, Protocols, and Guidance on Roles and Intervention*", funded by a grant from The California Endowment. Elizabeth has contributed to the development and/or implementation of several interpreter training programs such as the "*40-hour Healthcare Interpreter Training*" course currently in use at L.A. Care Health Plan, the "*Connecting World Central Valley Version*" program, and

the “*Spanish Bilingual Assistant Medical Interpreter Training*” created by Phoenix Children’s Hospital. To support health care providers, she also designs and conducts training for providers such as the cultural competency program for the resident nurses at CHLA. Her 4-hour training for providers on “*How to work more effectively with interpreters*” previously created for L.A. Care recently earned the health plan the “*Honorable Mention Award for Cultural and Linguistics Competency Excellence in a CME Activity*” from the Institute for Medical Quality.

Cynthia E. Roat

Cindy Roat is a consultant and trainer on issues related to language access in health care. She started working as a medical interpreter in 1992, after earning her MPH in International Health from the University of Washington. Ms. Roat is certified by the Washington State Department of Social and Health Services for both medical and social service interpreting. Her interest in systems change led her into training interpreters and interpreter instructors, then into training providers to work with interpreters, and finally into working with administrators on improving language access programs. Over the past decade, Ms. Roat has worked with large and small public and private institutions, in urban and rural areas all over the United States, making significant contributions in the areas of training, program development, policy formulation, advocacy and organizational outreach. She is the primary developer of *Bridging the Gap*, the country’s most widely-offered training for health care interpreters, as well as being the author of a wide array of key resources for the field. An engaging speaker, Ms. Roat is also in high demand as a conference presenter and trainer. Ms. Roat is a founding member and former Co-chair of the National Council on Interpreting in Health Care (NCIHC), as well as being known nationally as an energetic advocate for the field of health care interpreting and for language access in general.

Karin Ruschke

Karin is founder and president of International Language Services, a full-service agency providing on-site and telephone interpretation, written translation services, and training to clients nationwide. Ms. Ruschke received her M.A. from the School of Translation at the Monterey Institute of International Studies in California. Karin Ruschke has dedicated her career to bridging language and cultural differences in the health care setting.

In addition to running a community-based interpreting agency in Chicago, Ms. Ruschke provides consultation on effective set-up and implementation of language interpreting services to a variety of clients such as health care facilities, government agencies and community organizations. She has developed a comprehensive training program for interpreters in health care, which addresses the rigorous standards of accuracy, confidentiality, role and cultural-sensitivity, improving the effectiveness and professionalism of interpreting. She also trains providers in how to work effectively with interpreters. Ms. Ruschke is actively involved in all aspects of developing the medical interpreting industry and played an integral role in raising the awareness of standards for

medical interpreters. Karin was a technical advisor and lead researcher for the research project Hospitals, Language and Culture: A Snapshot of the Nation, sponsored by the Joint Commission on Accreditation of Healthcare Organizations in 2005. In addition, Ms. Ruschke was a member of the expert advisory panel of the American Medical Association on a project to assess ways in which physicians can overcome barriers to language access of their limited English speaking patients.

Laurie Swabey

Dr. Swabey has been an interpreter educator for over 20 years. She has been a nationally certified ASL/English interpreter since 1977 and has interpreted in a variety of community settings. Before moving to Minnesota in 1990, she was the director of the interpreting program at the University of New Hampshire, where she developed the Bachelor of Science degree in Sign Language Interpreting. Since moving to Minnesota she has developed and taught interpreting courses at the University of Minnesota for the Program in Translation and Interpreting. She currently is an Associate Professor at the College of St. Catherine, St. Paul, Chair of the ASL Interpreting Department there as well as Director of the CATIE center (Collaborative for the Advancement of Teaching Interpreting Excellence). She presents workshops on interpreting at the local and national level.

Jean Turner

Professor Turner taught ESL/EFL for a variety of schools beginning in 1976 with her Peace Corps service in Morocco. Before coming to the Monterey Institute she was Assistant Director of ESL Services at UCLA and taught classes on assessment at California State University, Los Angeles. She joined the Monterey Institute faculty in 1990 and in 1997 was the recipient of the Institutes' Dean's Award of Teaching Excellence.

She has consulted on test development projects and teacher training in the areas of assessment and curriculum design for various organizations, including Cambridge University Press, the Egyptian Fulbright Foundation, Educational Testing Service, AMIDEAST and Second Language Testing, Inc. Her current research interest in the areas of assessing advanced language, including assessing the skills and language proficiency of court and medical interpreters.

She is the author and co-author of articles published in the *Annual Review of Applied Linguistics*, *TESOL Quarterly*, *The Content-Based Classroom*, *CATESOL Journal*, *Preparing the Professoreiate of Tomorrow to Teach*, and *Language Testing*.

Education

PhD, Applied Linguistics; MA, TESOL, University of California, Los Angeles

STATE REPRESENTATIVES

Enrica J. Ardemagni - IN

Enrica holds a Ph.D. in Spanish from the University of Wisconsin-Madison, and Certificates in Cultural Competency, Medical Interpreting and Legal Interpreting. She is currently Associate Professor of Spanish and Director of the Certificate in Translation Studies at Indiana University Purdue University Indianapolis. Professor Ardemagni has published articles in national and international journals, book chapters on translation in medieval Spanish and Catalan literary texts, and edited eight medieval medical manuscripts. She teaches Translation Studies, Business, Medical and Spanish for Law Enforcement, and Medical Interpreting. She co-writes curriculum and teaches Spanish in the IU School of Medicine bilingual medical program. Professor Ardemagni is frequently invited to give workshops and presentations on cultural competency, translation and interpreting. She was one of the founding members of the Midwest Association of Translators and Interpreters (MATI), an Affiliate Chapter of the American Translators Association (ATA) and is the President of the MATI board. Professor Ardemagni also serves as the Administrative Chair of the ATA Literary Division, she is the Interim Co-Chair of the Organizational Development Committee of NCIHC, and the Chair of the Indiana Commission of Health Care Interpreters and Translators.

Carol Berg – MN

Carol Berg received her RN degree in 1977 from Lutheran Deaconess Hospital and a BS degree in Nursing in 1982 from the University of Minnesota. She continued her studies at the University of London and received a Master of Science in Community Health in 1983. She spent 10 years in Madagascar and worked as Director of Nursing Education and Consultant for the Primary Health Care Program of the Malagasy Lutheran Church. After returning to the U.S., she was employed as Assistant Coordinator of Community Health Services at the Minneapolis Health Department for two years which included supervision of interpreters. She then served as State Refugee Health Coordinator for 7 years at the Minnesota Department of Health.

Carol Berg is currently the Community and Public Health Manager for UCare Minnesota which is an HMO serving diverse members throughout the state. Carol has been involved in addressing interpreter training needs since 1989. She is the Chair of the Interpreting Stakeholder Group.

Maria Michalczyk - OR

Maria is the founder, director and instructor of the Healthcare Interpreter Training program at Portland Community College in Portland, Oregon, which was created in 1998. Her educational repertoire includes a B.S. in General Science, an M.A. in Anthropology from Portland State University and a degree in Nursing from the University of Guam.

Maria has worked as a healthcare professional for well over 30 years including critical care, utilization review services and general nursing duties. For a decade, Ms. Michalczyk worked for the Oregon Health Sciences University as a Medical Interpreter Manager and Diversity Training Instructor. Before taking this position she worked as an RN in a variety of clinical settings including working in Riyadh, Saudi Arabia. During the mid 1990's she was involved in the ASTM National Workgroup on setting national guidelines on medical interpretation. During 2000 she was the contributor to the Oregon Governor's Racial and Ethnic Health Task Force Report where she recommended the entire language found in the state wide report for medical interpreter services. Ms. Michalczyk was the major contributor and energy behind SB 790 to certify healthcare interpreters in Oregon. She testified for the bill and presented a speech to Governor Kitzhaber for the celebration of the bill becoming law in August 2001.

Ms. Michalczyk has served on the Board of Directors of the National Council on Interpretation in Health Care since 2001, serving as Co-Chair of the Organizational Development Committee and Co-chair of the Board and is presently Co-Chair for the Outreach Committee. Ms. Michalczyk has just completed a four year term as being the chair for the Governor's Council on Healthcare Interpreting in Oregon. Since February 2006 Ms. Michalczyk has been serving as the honorary Chair for the Japanese Association for Healthcare Interpreting. In April of 2006 Maria was awarded the "Governor John Kitzhaber Public Health Leadership Award" granted by Multnomah County Health Department in Oregon.

Armando Villareal - IA (not available)

Mauro Yanez – OK (not available)