

Lessons Affirmed and Learned
<p>(The MMIA Medical Interpreting Assessment for Certification Pilot, Forum on Meaning and Accuracy in Meaning, Forums on Certification)</p> <p>Prepared for the Expert Panel on Community Interpreter Testing and Certification</p> <p>Maria-Paz Beltran Avery, Ph.D.</p>




Development of the Prototype Instrument
<ul style="list-style-type: none"> ➤ The MMIA Certification Committee: an outgrowth of the Standards Committee that created and published the MMIA Standards of Practice in collaboration with EDC. ➤ Based on the Standards of Practice that were developed through a DACUM process ➤ Prototype developed in Spanish ➤ Spanish prototype pre-piloted in January 2001 ➤ Revised version piloted in 2003 with funding from OMH received through the NCIHC <p style="text-align: right;"><small>Maria-Paz Beltran Avery, PhD, 2007.</small></p>

Development (continued)
<ul style="list-style-type: none"> ➤ Funding supported a collaboration with the California Healthcare Interpreters Association (CHIA) ➤ Administrator/scorers were trained in both Massachusetts (10) and California (13) ➤ All administrator/scorers except one were native Spanish speakers (Caribbean, Central and South America, Spain) with minimum of 3 years experience as paid interpreters <p style="text-align: right;"><small>Maria-Paz Beltran Avery, PhD, 2007.</small></p>


Development (continued)
<ul style="list-style-type: none"> ➤ All administrator/scorers participated in a two day training ➤ Instrument was administered to interpreters in both Massachusetts (37) and California (46) over two days ➤ First day: written and sentence conversion (potential screening section) ➤ Second day: integrated skills (role plays) <p style="text-align: right;"><small>Maria-Paz Beltran Avery, PhD, 2007.</small></p>


Development (continued)
<ul style="list-style-type: none"> ➤ Final analysis conducted on 20 test results from Massachusetts and 22 from California (total 83); the rest were eliminated because of incomplete data, inconsistent administration ➤ Focus of analysis was on sentence conversion and role plays ➤ Report on pilot is available on NCIHC website <p style="text-align: right;"><small>Maria-Paz Beltran Avery, PhD, 2007.</small></p>

Purposes of the pilot
<ul style="list-style-type: none"> 🔍 To assess the validity of the prototype test: Does it measure what it is intended to measure? 🔍 To assess the reliability of the prototype test: Are the results of the testing consistent across administrations and scorers? 🔍 To try out different methodologies 🔍 To determine whether any of the modules can be used as screeners prior to administering the integrated skills module (role-plays) <p style="text-align: right;"><small>Maria-Paz Beltran Avery, PhD, 2007.</small></p>

	Purposes of the Medical Interpreting Assessment for Certification
	<ul style="list-style-type: none">  To determine basic entry level proficiency in medical interpreting  To provide a standard of quality  To provide interpreters with an assessment of their proficiency <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>





	Principles
	<p>1. Clear and public content standards</p> <p>Content standards define what an interpreter should know and be able to do.</p> <p>Defined in the MMIA Standards of Practice.</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>


	Principles
	<p> Clear and public performance standards</p> <p>Performance standards define the criteria for what counts as good performance.</p> <p>Defined in the MMIA Standards of Practice and in the performance rubrics.</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Principles
	<p> Uses authentic assessment methodologies</p> <ul style="list-style-type: none"> ✓ Authentic assessments are based on meaningful tasks require in the performance of the “job” that are related to the outcomes identified as essential. ✓ An authentic test not only reveals the level of achievement of the test taker to the examiner but also reveals to the test taker the actual challenges and standards of the field. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Principles
	<p>4. Addresses equity</p> <ul style="list-style-type: none"> ✓ Resources to acquire the knowledge and skills must be available to the candidate (i.e., training and education) ✓ The methodology and formats used in the instrument are accessible to all and comparable across languages while accommodating cultural and linguistic differences. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>


	Principles
	<p>5. Attends to consequential validity</p> <ul style="list-style-type: none"> ✓ Consequential validity refers to the accuracy of the decisions made on the basis of the instrument. ✓ The process and instrument are designed with a concern for the social consequences of the measurement. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Sections of the Instrument
	<ul style="list-style-type: none">  Basic human anatomy and medical terminology/vocabulary  Message conversions  Ethical and cultural issues  Integrated interpreting skills <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Section 1
	<ul style="list-style-type: none">  Knowledge of basic human anatomy and medical terminology/vocabulary <ul style="list-style-type: none"> o Modality: Paper and pencil o Methodology: <ul style="list-style-type: none"> o Diagrams and equivalencies o Matching terms with definitions o Scoring: Right or wrong <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Section 2
	<p>2. Understanding of ethical and cultural issues and the ability to make informed judgments based on such knowledge</p> <ul style="list-style-type: none"> o Modality: Written (Possible oral option) o Methodologies: <ul style="list-style-type: none"> o Scenarios: Multiple Choice o Scenarios: Open-ended questions o Scoring: rubric <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Section 3
	<p>3. Ability to convert oral messages accurately and completely</p> <ul style="list-style-type: none"> o Modality: Oral o Methodology: <ul style="list-style-type: none"> o Audiotaped messages from short and simple to longer and more complex (English to L2; L2 to English) o Scoring based on accuracy and completeness of units of meaning o Scoring: same as for Role Plays <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Section 4
	<ul style="list-style-type: none">  Ability to integrate the skill of oral message conversion and other interpreting skills <ul style="list-style-type: none"> o Modality: Oral, performance-based o Methodology: Role play simulation o Scoring: <ul style="list-style-type: none"> o Accuracy and completeness Performance rubrics <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Section 4
	<ul style="list-style-type: none"> o Scoring for Accuracy and completeness: 3 types of errors <ol style="list-style-type: none"> 1. Mistake: the meaning of the converted message is different than the meaning of the original 2. Omission: when a unit of meaning found in the original message is not found in the converted message 3. Addition: when a new unit of meaning is introduced into the converted message <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	So . . .
	<p>What did we find?</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Content/Face Validity
	<p>Based on a DACUM analysis of medical interpreting tasks</p> <p>Presentations to interpreter groups at national and local meetings indicated that the content included in the prototype instrument are appropriate and reflect the knowledge and skills that an entry level interpreter should have.</p> <p>Feedback questionnaires from test takers indicated that they found the content appropriate and fair.</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Predictive Validity
	<p>Statistical analysis of the correlation between performance on the sentence conversion portions of the test and performance on the role plays indicate that the Spanish to English conversion is a good predictor of role play performance (Pearson r of .744) but the English to Spanish is not.</p> <p>There is no available data currently to test whether performance on the test is predictive of performance “on the job.”</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Consequential Validity
	<p>The pilot did not provide data to determine whether the test met the criterion of consequential validity.</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Reliability
	<ul style="list-style-type: none"> o Consistency of the measurement across like respondents, across administrations of the instrument (stability), and across coders (scorers). o Statistical analysis focused on inter-coder reliability for the sentence conversions and role plays. o Each test was scored by two independent scorers. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Inter-coder Reliability
	<ul style="list-style-type: none"> o While inter-coder agreement, in general is reasonable, when you look into the degree of agreement between specific pairs, there is great variability. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

Inter-coder Reliability Sentence Conversions
<p>Agreement on each unit of meaning</p> <ul style="list-style-type: none"> ✓ 62% of the coder pairs had inter-coder reliability of .80 or higher in the English to Spanish Sentence Conversion ✓ 86% had inter-coder reliability of .80 or higher in the Spanish to English Sentence Conversion. ✓ A t-test showed that Massachusetts coders had significantly better inter-coder agreement than California coders. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

Possible explanation for MA scorer performance
<ul style="list-style-type: none"> ■ Instrument is based on MMIA Standards of Practice ■ MA administrators/scorers have been using MMIA Standards in training, supervision, etc. over many years ■ There has been more on-going communication, sharing of approaches among the MA administrators even prior to this pilot ■ We continued training for scorers beyond the original two days <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

How did test takers perform?					
<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top;"> <ul style="list-style-type: none"> ■ Based on an average of 80% accuracy <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; font-size: small;">Massachusetts</td> <td style="text-align: center; font-size: small;">California</td> </tr> <tr> <td style="font-size: x-small;"> <ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) </td> <td style="font-size: x-small;"> <ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%) </td> </tr> </table> </td> </tr> </table> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>	<ul style="list-style-type: none"> ■ Based on an average of 80% accuracy <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; font-size: small;">Massachusetts</td> <td style="text-align: center; font-size: small;">California</td> </tr> <tr> <td style="font-size: x-small;"> <ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) </td> <td style="font-size: x-small;"> <ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%) </td> </tr> </table>	Massachusetts	California	<ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) 	<ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%)
<ul style="list-style-type: none"> ■ Based on an average of 80% accuracy <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; font-size: small;">Massachusetts</td> <td style="text-align: center; font-size: small;">California</td> </tr> <tr> <td style="font-size: x-small;"> <ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) </td> <td style="font-size: x-small;"> <ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%) </td> </tr> </table>	Massachusetts	California	<ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) 	<ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%) 	
Massachusetts	California				
<ul style="list-style-type: none"> ✓ 6 passed SCES (30%) ✓ 9 passed SCSE (45%) ✓ 12 passed RP1 (60%) ✓ 10 passed RP2 (50%) ✓ 5 passed all SCs/RPs (25%) ✓ 5 passed RPs, failed SCs (25%) ✓ 5 failed all SCs/RPs (25%) 	<ul style="list-style-type: none"> 0 passed SCES (0%) 7 passed SCSE (32%) 11 passed RP 1 (50%) 10 passed RP2 (45%) 0 passed all SCs/RPs (0%) 9 passed RPs, failed SCs (41%) 7 failed all SCs/RPs (32%) 				

Next Steps for MMIA What we need
<ul style="list-style-type: none"> □ Development of a Blueprint □ Development of item bank for each module based on specifications □ Test the Blueprint with other languages and document modifications as needed □ A training and screening program for administrators □ A training and screening program for scorers <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have affirmed
<ul style="list-style-type: none"> ■ Measuring the skill of spoken language interpretating is a complex endeavor. <ul style="list-style-type: none"> – Measurement of language proficiency is itself a complex endeavor – Measuring the act of interpretation is different than measuring language proficiency; accuracy of the conversion and “untranslatable” terms are significant issues – The number of languages and the differences in the structures of the languages is staggering but most be addressed <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have affirmed
<ul style="list-style-type: none"> ■ Excellent interpreters do not necessarily make excellent administrators or scorers – these functions require different skills <ul style="list-style-type: none"> – Ensuring consistency of administration and scoring, especially when one of the central methodologies is a simulation, is not an easy task and requires extensive training, screening, and vigilance. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I have affirmed
	<ul style="list-style-type: none"> ■ 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, ways of knowing and expressing that knowledge, let alone a process that is valid and reliable across many administrations and versions. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I have affirmed
	<ul style="list-style-type: none"> ■ We, in the field of health care interpreting, have yet to come to a common understanding of what we mean by certification <ul style="list-style-type: none"> – We need to be clear about what it is and what it is not. – We need to go beyond describing what other’s call certification and make some judgments as to what are assessments/processes that meet the criteria for certification and which do not. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I have affirmed
	<ul style="list-style-type: none"> ■ As we develop rigorous, valid and reliable assessments in a limited number of languages, we need to find other rigorous ways of measuring and acknowledging the competencies of interpreters of other languages. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I have affirmed
	<ul style="list-style-type: none"> ■ There are a lot of questions about which we have little empirical data <ul style="list-style-type: none"> – Do we know what the minimum level of language proficiency is in order to achieve accuracy and intelligibility? – We say we want accuracy and completeness in the conversions but what do we really mean by this? Is there a hierarchy of importance in the concepts – e.g., are there concepts which must be in the conversion to qualify for accuracy while others may be omitted? <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I have affirmed
	<ul style="list-style-type: none"> ■ There are a lot of questions about which we have little empirical data (continued) <ul style="list-style-type: none"> – How do we measure accuracy of meaning? What does it mean, anyway? Do we have agreement on what we are measuring when we say “accuracy in meaning?” – What do we mean by “equivalence in meaning?” – Can there be accuracy without completeness? – How do we / should we establish cut scores? Is there a “good enough” score and, if so, how do we know it is good enough? <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Shared aspects of meaning:
	<ul style="list-style-type: none"> ■ Meaning has many levels/aspects <ul style="list-style-type: none"> – The content (what are the propositions contained in the message; the semantic content) – The context/purpose (what is the setting; what are other factors that enter into sense-making – e.g., culture, education, background, etc.) – The intent (what is it that the speaker wants to convey) <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

Key element: equivalence of meaning in the interpretation
<ul style="list-style-type: none"> ■ Equivalence of meaning between the message in the source language and the target language <ul style="list-style-type: none"> – What is the measure of equivalence? Some common descriptors <ul style="list-style-type: none"> ■ Did the message get through? Is it likely to be understood? ■ Is there clarity of expression? Is it intelligible? ■ The message presented in one language elicits in the listener the same image, message, connotations as <i>in the mind of the speaker</i> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

Key element: equivalence of meaning in the interpretation
<p>OR RATHER</p> <ul style="list-style-type: none"> ■ is equivalence of meaning <i>FIDELITY TO THE SOURCE MESSAGE</i> <ul style="list-style-type: none"> – Equivalence is in the conversion not in the understanding, i.e., fidelity of the conversion – This means that the receiver of the conversion responds to the message as if it had been heard in the source language <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have learned
<ul style="list-style-type: none"> ■ If we are to move to national certification, there is a lot of work we need to do in arriving at a shared understanding of what we are measuring, how we are measuring it, and what it means. ■ And then, we need to be able to convey all that clearly and explicitly to those who will be seeking certification and to those who will be availing themselves of the services of certified medical interpreters. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have learned
<ul style="list-style-type: none"> ■ In addition to developing an assessment for certification that is valid and reliable, there needs to be agreement on the purpose of the certification among the stakeholders. We cannot assume that we all share the same purpose. <ul style="list-style-type: none"> – Is it to determine competency in entry level skills and knowledge or is it to determine competency in many related skills? – Is there agreement on what the skills are and on what they look like in practice? <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have learned
<ul style="list-style-type: none"> ■ The certification process is not just about the assessment itself <ul style="list-style-type: none"> – it is also about the requirements to qualify to take the assessment; – this can often be a hidden but major point of disagreement among stakeholders based on differing beliefs about things like formal education, status, literacy, etc.; – Certification is a political process as much as it is a technical process. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

What I have learned
<ul style="list-style-type: none"> ■ Training on ethics in general and the profession's code of ethics is highly inadequate <ul style="list-style-type: none"> – Many interpreters do not understand the concept of ethics – While interpreters may be able to give a 'correct' answer to an ethical dilemma, they are unable to connect their answer to the profession's code of ethics and/or explain why the 'correct' answer is the ethical response <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I recommend
	<ul style="list-style-type: none"> ■ On-going, focused dialogue that builds knowledge <ul style="list-style-type: none"> – About what we know and don't know – About the mistakes we have made – About developing a common understanding of what competent interpreting looks like at different levels <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I recommend
	<ul style="list-style-type: none"> ■ Ongoing research <ul style="list-style-type: none"> – Current certification efforts should include a data gathering and research component that is 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. <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I recommend
	<ul style="list-style-type: none"> ■ Ongoing public awareness <ul style="list-style-type: none"> – With key stakeholders – About what certification is and is not – About what its benefits and limitations are – The rigor in development that it entails – The importance of transparency about the skills, knowledge, and levels of competency required – The need for opportunities to be prepared to succeed <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	What I recommend
	<ul style="list-style-type: none"> ■ Rigorous training - education <ul style="list-style-type: none"> – Of trainers/educators – Of interpreters – Not just the “how-to’s” (mechanics) but also the “why-to’s” (theory, reasoning, judgments) – With a common core of content – knowledge and skills – Based on standards of excellence for training and educational programs <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>

	Thank you!
	<p>Maria-Paz Beltran Avery, Ph.D. mavery@edc.org</p> <p style="text-align: right; font-size: small;">Maria-Paz Beltran Avery, PhD, 2007.</p>