Medical Interpreter Position Qualification (MIPQ)
Version 1.0.1

TEST DEVELOPMENT AND SPECIFICATION
INFORMATION FOR CUSTOMERS

Introduction:

When medical providers attend to patients without a common language of reference, an interpreter must serve as a conduit to facilitate the interaction. The standards of accuracy, preciseness, completeness and ease of communication required at this level should be commensurate with the consequence of the services being provided. Therefore, it is incumbent upon the organizations overseeing the provision of healthcare to communities and individuals requiring interpretation services to ensure that minimum standards are in place to confirm the ability of an interpreter to adequately perform without endangering patient safety or quality of care.

Encompassed within those requirements for accuracy, preciseness, completeness and ease of communication are certain pre-requisite skills, chiefly among them, the ability to speak and comprehend two or more languages fluently. In addition, the candidate must be able to accurately capture, retain, convert and transmit complex and nuanced medical concepts and instructions that may require a comprehensive understanding of human anatomy, physiology, pathology and pharmacology, as well as corresponding specialized medical vocabulary in both languages. Candidates must adhere to strict ethical guidelines and industry protocols to maintain transparency of interaction and allow for patient self-determination, with the primary goal of allowing patients to receive the same quality of care that would be provided to a patient without an interpreter.

The purpose of the Medical Interpreter Position Qualification 1.0 is to capture an example of a candidate's performance in a controlled yet realistic situation, and compare it to standards of performance set by Subject Matter Experts (SMEs) in the field to ensure a minimum level of competency required (pass/fail), as well as a reference score to allow for individuals and organizations to easily compare ability levels (A-F ranking).

Background:

In the past, many healthcare organizations have relied on weak proxies for interpreter ability, such as self-reported ability, or even ethnic background, when hiring and assigning interpreter duties. A step up from these flawed methodologies would be to conduct an informal interview (or partial interview) with the candidate in the target language, thus subjectively assessing basic

language fluency, but not necessarily evaluating the other skillsets required to fulfill the role. Some organizations may also use basic individual word matching/fill-in exercises to test for knowledge of terminology equivalents out of context, again only capturing partial skillsets.

In the last few decades as the profession and industry have matured, testing before hiring has become a standard practice in the medical industry, and more recently some states and independent non-profit groups have begun offering certification for medical/healthcare interpreters (in some common language combinations). It is important that the methodologies used to evaluate interpreters be structurally sound, with two primary requirements: 1. a solid objective statistical methodology, and 2. a complete and proportional review of each of the skills needed to perform the required work.

Professionally calibrated assessments designed by subject matter experts (SMEs) based on a JTA (job task analysis), ensure that relevant areas are assessed with applicable exercises (or highly correlating proxies). SMEs set certain standards for basic skills, set parameters for item difficulty, and determine relative weighting for different items and exercises. SME based assessments can be further improved by correlating item and exercise difficulty ratings with candidates' overall performance levels to measure the level of correlation between those difficulty levels and candidate outcomes (reliability coefficient) for both the aggregate and individual. This information in turn can be used to adjust relative weighting of those items and exercises within the tests. The 1.0.1 version of the MIPQ is an SME designed test.

In the future, we hope to further improve the predictive power of this assessment, directly correlating the weighting of individual items to candidates' long-term job performance rankings, quality assurance/peer review scores, and/or feedback from patients and providers.

Methodology and specifications:

The MIPQ will consist of a role-play scenario with a topic from general medicine/family practice/primary care/internal medicine. Candidates for general medicine interpreter will be evaluated on their performance within this scenario alone. Candidates for a specialized medical interpreter position will be administered an additional scenario from within the selected field (pediatrics, women's health, etc.), and will receive an independent ranking in both general medicine and the specialization.

The scenarios are designed with certain specifications in mind: They must include 40-60 medical terms, at least 20 of which are considered high level terms, and at least 20 of which will be considered low level terms. Low level terms are considered medical concepts with which the general population would be familiar, most direct cognates, and other terms deemed easy to interpret by a panel of subject matter experts. High level terms are considered medical terms which require an advanced study of and/or direct experience with high level medicine, as well as terms deemed difficult or moderately difficult by the SME panel. After an initial pilot, some terms may be moved between categories if difficulty and differentiation are higher than predicted. Overall

terminology performance is evaluated on a percentage basis for both high level and low level terms, which are then blended on a weighted scale to create an overall terminology score.

In addition to the terminology requirements, the scenarios must also fall within overall length specifications as measured by word count. Because the terminology is evaluated separately, the terms themselves are excluded from the word count when determining overall length, which can fall between 350 and 600 words.

In order to account for variations in length, accuracy errors will be proportionally weighted to the scenario's overall word count, thus maintaining a consistent relative difficulty level. Accuracy errors are sub-divided into high-impact and low-impact, based on the SME established guidelines previously referenced, and are assigned relative group weights within the overall score, those relative weights being overlaid onto the proportional weighting from the word count.

Assessments are administered by a human administrator, and the recording of the candidate's performance is reviewed and rated by a qualified, trained human tester. In addition to the overall percentage score ratings for accuracy and terminology, testers will also evaluate other aspects of performance, such as pace, professionalism, address and register, as well as grammar and syntax in each language.

Candidates receive scores for both terminology and accuracy, as well as an overall score. A score of 75% or higher is recommended in both accuracy and terminology to pass, even if a candidate's overall score is above 75%. However, hiring organizations may wish to set their own standards, depending on their own needs, demand for the language, and availability of resources.

Also available is an optional module on knowledge of medical interpreter ethics, protocol and industry standards (based on National Code of Ethics and Standards of Practice for Medical Interpreters published by the National Council on Interpreting in Healthcare in 2004 and 2005).

The written test on ethics, protocol and industry standards may be administered before the appointment to take the oral sections (scenarios). Most of the questions on the test evaluate complex understanding of the topics, and as a result, organizations may allow this section to be taken open book and/or without proctoring at their discretion. (However, although the oral portions are completed over the phone, LanguageStat highly recommends –but does not enforce- that organizations have candidates test in their offices or another proctored setting for the oral portion of the test.)

Organizations may wish to waive the optional module on ethics, protocol and industry standards if the candidate has other evidence of proficiency (a recently completed interpreter training course, for example).

Scenario development:

Scenarios are developed for either general medicine or a specialty by SMEs with extensive interpreting experience, guided by research of the conditions/ailments involved, and using terminology banks of specialty specific terms as well as general medicine terms. The scenario must

meet the overall length requirements (not including terms), as well as requirements of high level and low level terms, those levels being set by a survey of language specific SMEs.

After developing initial scenarios, a limited pilot was performed with candidates currently or formerly working in the medical interpreter field, as well as bilingual candidates working in related fields. Following the pilot, the terms lists were reviewed for overall difficulty level and individual correlation to overall performance, and some terms were reassigned on the basis or their correlation and coefficient. In addition, minor edits for flow and clarity were made before launching the 1.0 version. Further scenarios will be created to specification, and all scenarios may undergo periodic updates and edits as indicated based on cumulative data tracking of performance and item correlation.

All initial development will occur for the highest demand language (usually Spanish), and scenarios will subsequently undergo an additional level of review upon translation into other languages. Language specific SMEs will re-evaluate term difficulty level for each language pair. (For example: a low level cognate between English and Spanish may be considered a high-level term when testing English-Mandarin, where it is not a cognate). Scenarios may undergo language specific revisions and modifications to adjust for flow, logic and/or difficulty level if needed.

The section on ethics, protocol and industry standards is developed by SMEs based on the National Code of Ethics and Standards of Practice for Medical Interpreters published by the National Council on Interpreting in Healthcare in 2004 and 2005. There are a few initial questions about theoretical concepts relating to transparency, 1st person/3rd person speech, and consecutive/simultaneous modes of interpreting, followed by a series of hypothetical situations and examples designed to elicit complex comprehension of the applicable concept through their form and content, and a series of true/false questions. Introductions are not evaluated during the test, as scripts and protocols for interpreter introductions vary greatly depending on the employer or training organization.

Test Reporting: Please see the next page for a sample report.

anguage STAT

anguage STAT

MEDICAL INTERPRETER POSITION QUALIFICATION (MIPQ)
-ENGLISH/SPANISH-
General Medical Interpreter

Candidate Name: *Jane Doe*
Test Date: *January 12, 2015*

Terminology Score: 90%
Accuracy Score: 86%

PASS **B+ grade (88.0% overall score)**

75% or higher in both terminology and accuracy required to pass.

Cultural competency component: Not included
Industry standards, ethics & protocol: Not included

English	Spanish
Accent/Pronunciation: Native level inflection Grammar and Syntax: Proper grammar and syntax	Accent/Pronunciation: Mild accent, easy to understand Grammar and Syntax: Occasional minor errors that did not impede comprehension

Skills review

Pace: Interpreted without hesitation and requested a limited number of repetitions.
Professionalism: Maintained professional demeanor.
Vocabulary: Maintained the register of the original speech, and did not insert any words from either language while interpreting into the other.

Tester recommendations: We recommend that Ms. Doe review Spanish grammar and syntax, and continue speaking for further refinement of her language skills. We also recommend that she review medical terminology in both languages, and that she focus on note taking while interpreting, to continue improving the accuracy and completeness of her interpretation.

Grading scale	*A+ 98-100%* *A 93-97%* *A- 90-92%*	*B+ 87-89%* *B 83-86%* *B- 80-82%*	*C+ 77-79%* *C 73-76%* *C- 70-72%*	*D+ 67-69%* *D 63-66%* *D- 60-62%*	*F 0-59%*