

Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

[Clip 1: EIPA](#)

[Clip 2: EIPA, continued](#)

[Clip 3: Questions for Mentors](#)

Clip 1. EIPA

Hello, I'm Brenda Schick. I'm here to explain a system for evaluating educational interpreters. The name of that system is the Educational Interpreting Performance Assessment, abbreviated EIPA. It was originally developed by myself and Kevin Williams, who works at the Boys Town Research Hospital. EIPA is really a system, or procedure, for evaluating educational interpreters. The program is designed for any interpreter who works in a public school setting, from kindergarten all the way up through twelfth grade. So, the system focuses on interpreters who work with children, as opposed to adults. One unique thing about the EIPA is that it can be used to evaluate any kind of sign language or sign system, whether it be ASL, PSE, MCE, English based signs and so forth. The school and the interpreter can decide on what type of sign to use and we are able to evaluate them regardless of what they may choose. I think it will help to understand a little bit about the history of the EIPA, because it has really changed over time. The first EIPA was developed mostly as a means of providing feedback to interpreters in order to improve their skills. But it happened that at the same time, schools and states all over the country were looking for some way to evaluate educational interpreters. They asked us to help them and the EIPA then changed so that it became more of a diagnostic feedback system and increasingly states and school districts are using it for certification purposes.

The EIPA has several forms and processes because of those changes that have occurred over time. You might be wondering how exactly the EIPA works in terms of its procedure. How exactly is it done, in other words. In summary, there are three steps. The first requirement is collecting samples of an interpreter's work in the form of videotapes of both expressive and receptive work. So, signing within the classroom, as well as voicing the signing produced by a child. That sample will be used later for evaluation. The evaluation component uses a specific form with about forty-four different skill areas that are evaluated. Finally, the assessment requires feedback with the interpreter regarding exactly what will help them improve their skills. So, that's just a basic overview of the assessment; it requires a sample, an evaluation, and feedback.

Now, let's discuss collecting a sample of an interpreter's work. There are two different ways in which that can be done. One is 'live;' that is, actually being in the classroom. The second is what we call 'stimulus videotape' and we have those prepared ahead of time for the interpreter to watch. The live version is done by the interpreter selecting a day, or a lesson, in which they will record



These videos and materials were produced for the Master Mentor Program by Project TIEM. Online, Dept of Ed, Grant H160C00003, 2001-2002.

Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

themselves. Understand that the material must be challenging and complex, not something that is a review, for instance. The classroom teacher must sign off and approve that, in fact, the material is new and the interpreter has not had an opportunity to go over it and review it ahead of time. Of course, all interpreters do preparation work, but it's not to a degree of over-preparation. And then the interpreter simply videotapes themselves while they are interpreting within the classroom. And you can see the child, other classmates, the teacher and so forth. It is an actual, real-live situation. For the receptive portion, we use a complex interview technique. This involves the deaf child, a hearing person who does not know sign and the interpreter all sitting together. The hearing person asks the child complex questions; these are not simplistic questions like "What is your name?" or "What is your favorite color?" Rather, it is something complex and interesting. The interpreter, of course, interprets the dialogue, voicing the deaf student's commentary and signing the hearing person's commentary. The three of them are videotaped having this interaction. Both of those instances, the live classroom and interview, are used within the evaluation.

Now, the second method of collecting samples is to use pre-prepared, stimulus videotapes. These videotapes are very unique. We have different versions of the tapes. For instance, there is an elementary version and a secondary, or middle school, version. Each of these shows several different classroom situations. I think the elementary version has six different classrooms ranging from first through sixth grade. And there is about ten minutes of tape for each of those. Before the interpreting session, the interpreter receives a copy of the lesson plans so that they can visualize what will be discussed, some of the key concepts or vocabulary that will be used, and in general what the lesson will be about. Therefore, it does not just come out of the blue. Instead, they have some time to think about and consider what will be happening before they watch the videotape. Once they are watching, they are recorded providing an interpretation for these classroom situations. What's nice about this is that it is an actual, real-life classroom situation. So, there are children raising their hand to ask a question, or children who are chitchatting, and children teasing one another, and so forth. All of that is on the videotape, so that the interpreter is sometimes required to change register, for instance. The interpreter will have to take many things into account.

For the receptive portion, we also have videotapes, but they are of children signing. We'll have a child commenting on different topics and someone asking them questions about it. The interpreter has to receptively take in what the child is signing and then provide a voiced interpretation for that. And all of that is videotaped as well. So, while they both use some previously developed videotapes, they cover educational situations and children. The secondary version is almost the same as the first in that they share the same concept. It is in a classroom with a teacher being videotaped. The interpreter watches that and



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

provides a signed interpretation while being recorded on videotape. The receptive portion involves watching an older child being asked questions and signing a response, for which the interpreter is videotaped voicing the commentary. Both videotapes require the interpreter to provide sign-to-voice and voice-to-sign interpretation so that they can both be used in the evaluation.

Now, the live version that we discussed previously is still used quite a bit, but it has some problems associated with it. It is good for diagnostic purposes, but it is a little more problematic when used for certification. For example, the interpreter is self-selecting the lesson and everything that will be videotaped. And that means that the evaluation team has a difficult time predicting what will happen within that classroom. Or if the interpreter, for instance, never uses numbers because the teacher never uses them, how do we know that the interpreter is able to use them effectively? Those types of things definitely make it more difficult.

Our evaluation team always has one deaf person out of the three person team. One person is required to be deaf because we feel that often a deaf person will see the bigger picture. So, because the classroom situation is not captioned, there is no documentation of what the hearing person is actually voicing. Thus, it is hard for the deaf evaluator to fully participate, because they are dependent on the other two team members to fill in what the teacher is saying. That is not the most ideal situation. A third problem with this method is that the classrooms can vary so much in their levels of difficulty. Some of the teachers are speaking extremely quickly, making it difficult to follow. Another teacher might be much more laid back, creating an easier situation for that interpreter. So, using the live version does not ensure an equal level of complexity and difficulty, so that there is an equal playing field. In addition, sometimes an interpreter can prepare too much for the assessment. What I mean by that is that an interpreter can't use skills they don't have. They cannot invent new skills. But, maybe the teacher will give them more materials than is typical. Or the interpreter will have an opportunity to practice at home before hand. And even though it is rare, it does happen. Also, remember that we are having an interview with a real child and sometimes that can also be a problem with the live version.

Often the complexity and the quality of the child's language is actually a function of the interviewer. Perhaps the interviewer asks easy, simplistic questions such as "What is your favorite color?" or "What is your favorite sport?" With those very easy types of questions, not much depth is required in terms of receptive skill or providing an interpretation. It makes it a very easy situation for the interpreter and that can be a problem. Sometimes, the child is low-functioning, doesn't sign much, or signs minimally. For the interpreter, this lessens the requirements of the job because the child is not doing very much. Also, some of the



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

children in classrooms actually use some speech. The interpreter may have become dependent on this in terms of providing a voice interpretation, so that the voicing is not very difficult and doesn't present much of a challenge for the interpreter. And sometimes the interpreter has been with the child for a very long time, perhaps many years. They know that child very well and it makes it easy to voice for that particular child. If they are put in a situation in which they are voicing for a completely different student it would be a different story. So, the fact that they may be too familiar with that one particular child can also be problematic. And I've also seen it happen several times where the interpreter may be struggling with a voice interpretation, so the child adjusts their language to make it much simpler. They know to make it simpler because the interpreter can't handle a more rapid rate or advanced level of signing. The result is that it may make the interpreter look good in terms of the interpretation they provide, but in actuality it is due to an adjustment in the language complexity level. The live version, then, has several potential issues and problems. It is still very functional and useful and shows a lot about that interpreter's skills, but the videotaped version really controls the situation more.

Clip 2. EIPA, cont.

Now, we have a sample that is ready to be used for evaluation. The evaluation is done using a team of three evaluators, who are all trained using the EIPA method. It requires specific, lengthy, in-depth training in order to observe and rate the performance of an interpreter. Of the three evaluators, one is required to be deaf. Also, their language must match that of the interpreter. For instance, if the interpreter is supposed to be using Signed English, the two hearing evaluators need to be proficient in Signed English. The deaf person does not necessarily need to be proficient in Signed English because they are really functioning in a different role. They are looking for skills and domains that are universal to all sign languages and sign systems.

You will recall that I said the EIPA is often used in different states. When the state establishes the EIPA program, we go in and train people within the local area to be evaluators. The evaluators chosen must already have skills in interpreting prior to being selected, of course. They are trained and then the three team members come together and view and evaluate the videotape. One positive aspect about using a local team is that they are knowledgeable about what the signs look like in that particular area and any regional variations that may exist, and so forth. They are aware of all of that. If we were to send a videotape to evaluators in another state, those people might not be aware of nuances that are particular to a region. There is one other method of evaluation that is a little bit different. Kevin Williams, the co-developer of the EIPA, works for the Boys Town National Research Hospital. Boys Town really established a



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

diagnostic center for educational interpreting so that any state, any interpreter, any individual can ask for an evaluation and Boys Town would use their training team for the evaluation. And increasingly, interpreters are having their evaluation done by the Boys Town National Research Hospital Diagnostic Center. During the evaluation, the team uses a specially designed form, which has about forty to forty-five different skill areas. Some examples of skill areas are grammar, use of space, facial expression, eye gaze, fingerspelling, using numbers, and so on and so forth. The team will observe the videotape and looking at each one the areas, will score each skill area. All together, the evaluation takes more than an hour. The score itself involves six different levels. A zero would indicate that no skills whatsoever are present. A one (1) denotes a Beginner. Levels increase to Advanced Beginner (2), Intermediate (3), Advanced Intermediate (4), and finally Advanced (5). The evaluators observe for each area and tabulate them to determine and average for each area. You might be curious about what the results typically look like or the skill level of most interpreters.

I'll summarize for you the published data, which you have in the handouts distributed to you in the form of a table. We have data from fifty-nine (59) interpreters from all over the state of Colorado. The data was collected over three years time. Of those fifty-nine interpreters, more than one half of them were below a 3.5. The state of Colorado has established a score of 3.5 as the very basic minimum. Below that level, you are not eligible to interpret in the public schools. More than half of the educational interpreters, then, didn't have the minimal skills needed to interpret in the classroom. They are still out there interpreting, though. Also, we have data from Kansas, using a different evaluation team. The data from Kansas looked almost identical to that collected in Colorado. Again, more than half of the educational interpreters working in Kansas did not have the minimum skill level. You'll recall that the EIPA has two different versions: live and videotape. In Louisiana, they have done a smaller research study to determine if interpreters receive the same score using each version. They found that with the stimulus video there was almost always a score of one level below the score achieved in the live version. The correlation was 0.94, meaning that it really didn't matter if they used the live or videotaped version of the assessment. We had a valid measure of interpreting skills.

So the question arises, why would there be the discrepancy of one level lower when using the videotaped version. This is due to the fact that in the live version there are not a lot of controls. Any interpreter will perform better if they have some familiarity with the situation, if they know the classroom and the child. So, of course, we can expect that there would be a difference. But, the data shows that consistently the difference is one level. Also in Louisiana, we conducted a study of interpreters who used the elementary videotaped version and those who used the secondary videotaped version. There was a comparison



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

of the scores using those two versions, and we determined that it was almost exactly the same. It didn't matter if the interpreter was interpreting in an elementary setting or a secondary setting. Interpreting in an elementary classroom didn't mean it was easier or make it look like they were improving because they were interpreting at an easier level of material. Both settings are equally difficult because they require the interpreter to use a lot of complex skills regardless of which setting it was. So, the assumption that the scores from the elementary evaluation will always be higher than those from the secondary assessment is not true. They are in actuality the same.

Often, people want to know if we have any psychometric data to show that the EIPA is valid. We do have some data and we are still collecting more. We have same data for inter-rater reliability. To do that involves having two different teams, Team 1 and Team 2, and one interpreter's videotape. Both teams evaluated one interpreter and came up with similar determinations for that interpreter. This means that the assessment has good reliability. We also have data for "test-retest reliability." That involves one interpreter who is evaluated using videotapes and all the regular evaluation materials. At a later date, that interpreter is evaluated again using a different set of tapes and a different evaluation team. The question that is being addressed is whether or not the results will be consistent. Statistics show that the results for the first and the second testing are the same. This means that if you were to test a person one time and test them again later a second time, you could expect to see the same results. The understanding in all of this, the testing and retesting, the levels used, the language complexity involved, etc. are all the same. If you were to change the test from elementary to secondary level, you may see some differences. But, if the test and retest involve the same levels, the same language, the results will be the same. We are still collecting more psychometric data now to create a larger and ever increasing set of data.

Clip 3. Questions for Mentors

What does one do with an interpreter's EIPA when they have it at hand? What does one look for? What can be known from it? Remember that the EIPA has changed over time. There are several things that you should know when you are looking at that evaluation.

The first consideration is which version of the EIPA you are using. Remember, there are both live and stimulus-videotape versions. You will want to know which one you are working with. This is because, as you'll recall, the data from Louisiana showed us that interpreters who were evaluated using the live version vs. the stimulus-videotape version have a difference of one level in their results. So, for instance, it might be 3.0 with the stimulus-videotape or 4.0 with the live



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

version. Which version of the assessment has been used, then, is extremely important. Also, remember that the disparity between the levels of the two versions has a very high correlation of 0.94. This means that almost 100% of the time, the two different versions will produce a difference in score of one level for all interpreters.

The second question to ask is who is actually providing the supervision and overseeing the testing. Some states are very “informal” and the person responsible for the supervising the system in actuality knows nothing about educational interpreting. They are not familiar with the EIPA and function in more of an administrative role. Maybe the evaluation team is making it very easy and just passing everyone. The administrator who is supervising and overseeing all of this is not aware of it because they don’t know the assessment process. It’s really important, therefore, to ask and find out who oversaw the qualifications of those involved in the assessment. Some states are very strict and diligent and have very strong evaluation teams. On the other hand, some states, unfortunately, are very informal in how they go about it.

The third thing to consider is whether or not the results come from using the elementary or the high school version. Now, as you’ll recall, I explained earlier that it didn’t matter which whether it was the elementary or high school version that was used in determining the reliability of the score. The scores are almost identical regardless of which one is used. But, it’s important to understand that it can be different for different interpreters. We don’t know for a specific interpreter between elementary or secondary what their particular score would be. We don’t know if it would be the same or if it would be different. So, if you want an interpreter for a high school position, and you have the EIPA results from an elementary version, you may want to repeat the assessment in order to make an appropriate determination. You may want to evaluate them again. So, you’ll want to know which version they were assessed on when looking at their scores.

The fourth consideration is what was used in terms of the language production. Was it sign language or a signed system? Remember the EIPA can be used with any sign language or signed system: ASL, PSE, more of an MCE type system, etc.. It’s important for you to know which language or system they were using while they were being evaluated. If you want an interpreter for an ASL position and the interpreter’s assessment was done using MCE, that may be a situation where you will want to reevaluate. One can’t assume that having scored well using the MCE that it correlates to skills using ASL. That unfortunately is not possible.



Translation: Brenda Schick Presentation

translated by Kathryn Bennett, MA, CI

A fifth point of consideration is the interpreter's skill levels. Remember, the EIPA evaluation form has about forty-five different skill areas which it assesses. You'll want to look carefully at where their strengths and weaknesses lie. We've found that most of the interpreters will do better in the area of vocabulary, because they know a lot of vocabulary, or a lot of signs. And that's fine. But, there are also skills of facial expression, use of space, or grammar that may be lower in score. You might see a good score, but you'll want to look at whether there are discrepancies in particular areas, or if there is consistency across areas and the interpreter does not have any deviation in score between areas. Perhaps, the interpreter's average score is influenced by one particularly very strong skill area, such as vocabulary for instance. But, all the other areas they consistently score very poorly. You'll want to look at all of those skill areas together, not just the average score for the EIPA.

The sixth consideration is something you'll need to decide for yourself, or the school district, or what have you. This entails what level is considered successful or acceptable. You might recall I said Colorado accepts 3.5 as the very bare, basic minimum, and anyone who scores below that is not considered a candidate. It's important to understand that 3.5 is not considered a proficient interpreter. A proficient interpreter would be at 5.0, the Advanced level. But, school districts often will hire someone who is not at that proficient level. We want to hire someone with good, solid skills, not who is mediocre. We want them to have good skill, even though they may not be at the very top in their field. Each school district will have to make that decision about what that minimum standard of acceptance is. That's a difficult discussion, because all of us, of course, would prefer to have every interpreter be the best. We want everyone we hire to be a 5.0, advanced interpreter. But, it's a fact that often school districts can't find someone who is at that advanced level and they'll end up waiving the requirements and hiring someone who scored a 2.0. There needs to be a compromise. If having every interpreter be a 5.0 is not realistic, then a 3.0 or a 3.5 is an acceptable standard. What level you or your school district will come to determine as the standard is not something we can tell you. It's something that needs to be reached individually.

