ABSTRACT: The Sign Language Proficiency Interview (SLPI) is an interview tool that is widely used across the United States and Canada for assessing a person’s sign language skills. In the traditional SLPI 3-rater team procedure, 3 raters independently rate a video recording of an interviewee’s sign language interview and subsequently discuss and negotiate an official performance rating by reference to an explicit construct of scaled sign language skills. This construct defines each of 11 scale levels from low beginning sign language skills (No Functional Skills) to native-like sign language skills (Superior Plus). Appendix A provides descriptors of the functional language use construct that is associated with each SLPI rating scale level.

The SLPI was originally developed and piloted at the National Technical Institute for the Deaf (NTID) for use in evaluating the sign language skills of NTID faculty and staff (Newell et al., 1983), with the first formal implementation of the SLPI occurring for residential staff at the Louisiana School for the Deaf (LSD) (Caccamise, Newell, & Mitchell-Caccamise, 1983). Since that time, the SLPI has come to be widely used across the United States and Canada, having been implemented at more than 60 academic and vocational rehabilitation programs. Despite its wide use, there are no published reports on the reliability or validity of SLPI ratings. Given its importance to hiring
practices, tenure and promotion, and program admission and graduation, such studies are necessary to help ensure that the ratings that are received by SLPI interviewees are consistent and accurately reflect their skills. In addition, information obtained from such studies may assist in planning and providing in-service training for SLPI team members.

There are three options in use at various SLPI evaluation sites for obtaining SLPI ratings from this interview. These options are the traditional SLPI 3-rater team procedure, the SLPI 2-rater team procedure, and the SLPI individual rater procedure (Caccamise & Newell, 2007a). This article is concerned with the traditional SLPI 3-rater team procedure. In this procedure, the interviewee is recorded and this video recording is then rated independently by three trained SLPI raters. The three raters may conditionally provide a second set of independent ratings if all three raters’ initial ratings are not within one rating level of one another. Subsequent to these independent ratings, the three raters convene to discuss the interviewee’s language behaviors, with explicit reference to the SLPI rating scale construct, and they attempt to negotiate an official rating that is consistent with that construct.

In the present study, we evaluated the interrater reliability of raters’ independent ratings before negotiation. In addition, we examined the degree of deviation of independent rater’s prenegotiation ratings from their postnegotiation official ratings. A rater’s prenegotiation rating is affected by the SLPI construct as well as by raters’ idiosyncratic perceptions of an interviewee’s behaviors that are irrelevant to the SLPI construct. Deviations of raters’ prenegotiation ratings from their postnegotiation ratings reflect how much the construct validity of the SLPI depends on the negotiation process.

**Interrater Reliability**

Interrater reliability in this study refers to how consistently the scale ratings of the three raters within an SLPI team agree before discussion and negotiation. We obtained SLPI results from a sample of NTID faculty, staff, and students, and evaluated interrater reliability by determining the total relative proportion of individual rater’s ratings before negotiation that were within one or more scale units of their team members’ individual ratings.

**Construct Validity of Prenegotiation Ratings**

Construct validity, in the context of language performance, refers to the notion that a test should accurately reflect the principles of a valid theory of language use and learning. In this respect, a good construct-referenced test of language performance should include features of actual language use, including (a) interactions among people, (b) accommodating unpredictability in the reactions of language users, (c) language use in context, (d) allowing immediate clarification of the purpose of a person’s language behaviors, (e) adjusting to the linguistic level of a person while maintaining authentic language, and (f) being performance based in a way that the quality of a person’s language is revealed (Finocchiaro & Brumfit, 1983). As Moss (1992) pointed out, construct validity is central to the validation of testing methods and practices for performance assessment.

The SLPI is a construct-referenced test that incorporates the above features of language use (see Appendix A for construct descriptors). The face-to-face format of the SLPI provides a natural context for these features to emerge. As Clark (1978) pointed out, from the standpoint of validity, tests that involve face-to-face interviews approximate real-life communication as closely as possible in a testing environment.

SLPI raters are trained to respect the principles and practices of SLPI administration based on the SLPI construct. Therefore, raters’ prenegotiated, independent ratings are construct referenced in the sense that they consciously attempt to adjust their judgments to an explicit description of the SLPI construct. However, it is well known that the accuracy of judgments of expert raters is influenced by characteristics of the individual raters themselves, including situational factors; health; idiosyncratic biases; level of experience in the specific judgment task; and other physical, psychological, and professional variables (Stewart, Roebber, & Bosart, 1997). Such variables can compromise the ability of raters to accurately apply their formal knowledge of the reference construct. Therefore, a variety of rater-specific (idiosyncratic) characteristics may introduce construct-irrelevant variance into SLPI raters’ ratings that may cause those ratings to drift away from the true level of interviewees’ sign language skill to different degrees across different interviewees.

In order to reduce these rater-specific sources of construct-irrelevant variance, the traditional SLPI 3-rater team procedure was designed to include a rater negotiation process in which the rating team members discuss and negotiate their official rating. During negotiated discussion and review of an SLPI, raters calibrate their common understanding of the characteristics and quality of an interviewee’s sign language expressive and receptive skills with reference to a standardized linguistic performance profile for signers at different levels of sign language proficiency. As such, each rater’s compliance with the explicit construct of scaled sign language skills that underlies the SLPI is reinforced within the immediate time frame of the evaluation process for each interviewee’s language sample. This in situ self-calibration with respect to a linguistic performance profile is consistent with Moss’ (1992) emphasis on incorporating principles of a valid theory of language use and learning into language testing. It therefore results in an official language skill rating that, in principle, provides greater construct validity than an individual rater’s rating or than an average of three raters’ independent ratings.

Caccamise • Samar: SLPI Reliability and Validity 37
In this study, we investigated the construct validity of raters’ prnegotiated independent ratings. Construct validity before negotiation was evaluated by determining the total relative proportion of these ratings that were within one scale unit of the negotiated official ratings. This comparison provides an intuitive estimate of the impact of the self-calibrating negotiation process on reducing specifically idiosyncratic sources of construct-irrelevant variance. For example, a small proportion of an individual rater’s ratings deviating from the negotiated official ratings would imply a small impact of idiosyncratic rater factors on the construct validity of their individual ratings.4

**METHOD**

**Participants**

Participants were all NTID SLPI team members (n = 34) who conducted SLPI ratings for NTID faculty, staff, and Master of Science in Secondary Education (MSSE) students who took the SLPI from the 2003–2004 academic year (AY) fall quarter through the 2005–2006 AY fall quarter. During this time, the 34 NTID SLPI team members provided ratings for 160 SLPIs. Of the 34 team members, 19 (55.9%) were deaf and 15 (44.1%) were hearing, with 16 (47.1%) working in NTID sign language and interpreter faculty/staff job positions and 18 (52.9%) working in other NTID faculty/staff job positions. All SLPI team members satisfied the criteria of an Advanced Plus-Superior Plus skill level range for participating in SLPI training, and all had successfully completed SLPI training conducted by NTID SLPI trainers. In addition to the initial training required to become an NTID SLPI team member, in-service training is provided annually on an individual and group basis for NTID SLPI team members. This training is planned by the NTID Office of Communication Assessment Services Coordinator based on monitoring of services provided by SLPI team members. By the end of this study, SLPI team members’ length of NTID SLPI team service ranged from 1 to 15 years, with 12 having 12–15 years of service, 15 having 4–9 years of service, and 7 having 1–2 years of service.

**Procedure**

**SLPI procedure.** All 160 SLPIs were conducted using the NTID standard interview procedure as described by Caccamise and Newell (2008a). In brief, all SLPIs involved a video-recorded one-to-one conversation between a trained SLPI interviewer and the person taking the SLPI. Interviews generally last approximately 20 min and cover three general topics: work/schooling, background, and hobbies. Specific questions for each general topic vary according to the interviewee’s responses.

**SLPI rating procedure.** Within 3 months of interview dates, these interviews were then rated using the standard SLPI 3-rater team procedure as described in Caccamise and Newell (2008b). This rating procedure is as follows:

- Three raters view each interviewee’s SLPI recording and, using SLPI Individual Rater Worksheet B (Appendix B), they independently describe and provide examples of the interviewee’s sign language skills. Based on a comparison of their descriptions of the interviewee’s sign language skills to the SLPI rating scale (Appendix A), raters then record their first independent ratings for the interviewee on their individual rater worksheets. A fourth person collects the three raters’ worksheets.

- If the first independent ratings of the three raters are within one rating level of each other, the individual rater worksheets are returned to the raters and they begin the discussion needed to complete the SLPI Rater Discussion Worksheet B (Appendix C). Following completion of the discussion worksheet, raters compare their agreed-to description of the interviewee to the SLPI rating scale and record their final independent ratings on their individual rater worksheets. If all three raters are in agreement, this rating is recorded as the official rating on the SLPI Rater Discussion Worksheet B. If all three raters are not in agreement, the raters may recommend that the interviewee be rated by another SLPI rating team or that the interviewee be scheduled for another SLPI.5

- If the first independent ratings of the three raters are not within one rating level of each other, the individual rater worksheets are returned to the raters, although the ratings are not shared with them, and the raters view the interview a second time.

- Following their second viewing of the interview, the raters record their second independent ratings on their individual rater worksheets. If all three raters are within one rating level of each other, they begin the discussion needed to complete the discussion worksheet and to decide their final independent ratings. If all of the raters are not within one rating level of each other, they discuss whether they believe the interview to be ratable. If the raters agree that the interview is ratable, they begin the discussion needed to complete the discussion worksheet and to decide their final independent ratings. If the raters believe

---

4The negotiation process does not reduce or eliminate all sources of rater-related construct-irrelevant variance. Biases and experiences that are common to all raters in a team will not generate discrepancies in initial independent ratings among team members and will therefore not be mitigated by negotiation. Such biases might include culturally entrenched attitudes or perceptions of psychologically salient but construct-irrelevant factors such as cultural standards of physical attractiveness, race and ethnicity, communication philosophy, and cultural identity. Therefore, the validity analysis in this article is limited to estimations of the influence of idiosyncratic rater characteristics on the construct validity of SLPI official ratings.

5Because the SLPI Advanced Plus, Superior, and Superior Plus ratings are reported to interviewees as a single official rating, the Advanced Plus-Superior Plus range, these three ratings are considered to be a single rating level for the NTID SLPI 3-rater team process; all raters must agree on the same exact rating level for all other ratings.
that the interview is not ratable, they record not ratable on their individual rater worksheets and on the SLPI Rater Discussion Worksheet B official rating line.6

- If the three raters are unable to agree on an official rating, or they believe that an interview is not ratable, the SLPI coordinator may inform the interviewee that the interview is not ratable and provide the interviewee with the option to schedule another SLPI or, if he or she believes that the interview is ratable, the coordinator may give the interview recording to a second rating team.

- If an interviewee, after reviewing his or her SLPI interview at an SLPI follow-up meeting, believes that he or she should have received higher ratings, the interviewee may request that the interview be rerated. Also, if an interviewee believes that he or she did not demonstrate his or her highest sign language skills, the interviewee may request another SLPI. Re-interviews and reratings are always conducted by SLPI team members who did not conduct the first interview and rating for an interviewee.

- When a rerating is requested, if the rating by the second team is in agreement with the rating by the first team, this rating is awarded as the interviewee’s official rating. If the first two rating teams are not in agreement, the interview recording is given to a third team. If the rating of the third team is in agreement with the rating of one of the first two rating teams, this rating is awarded as the interviewee’s official rating. If there is no agreement across the three rating teams, the interviewee is scheduled for another interview.

RESULTS

Of the 160 SLPIs that were conducted, 157 (98.1%) received SLPI official ratings. The rating teams for the remaining three interviews (1.9%) were unable to reach agreement; hence, these interviews could not be assigned official ratings. Only eight (5%) of the 160 interviews required reratings, six of which resulted in official ratings and two of which did not.

The SLPI protocol allows for the possibility that an interview might not be ratable and, therefore, might not result in an official rating, as happened with three of the interviews in this study. Interviews that do not meet the criteria needed for an official rating cannot affect the reliability of official ratings, nor can they contribute construct-irrelevant variance to official ratings. Therefore, the analyses below are based on the 157 interviews that entailed interviews that met the general SLPI criteria to receive official ratings.

Table 1 shows that the highest number of interviewees receiving official ratings achieved an official rating of Advanced Plus-Superior Plus range (43, or 27.4%), followed by Intermediate (31, or 19.7%), Intermediate Plus (29, or 18.5%), Survival Plus (21, or 13.4%), Advanced (15, or 9.6%), Survival (14, or 8.9%), Novice Plus (3, or 1.9%), and Novice (1, or 0.6%), with no interviewee receiving no functional skills as an official rating.

Interrater Reliability

Table 2 shows that 86.6% of the raters provided first independent ratings that were either the same as or within one rating level of those of the other members of their rating team, and 13.4% of the raters provided first independent ratings that were not within one rating level of those of the other members of their rating team. Most of these latter raters provided second independent ratings that were within one rating level of those of their rating team members. Specifically, after second independent ratings, 96.8% of the raters provided ratings that were either the same as or within one level of those of their rating team members, which is an improvement of 10.2%.

Table 3 provides an analysis of the first and second independent ratings grouped by SLPI rating scale levels. The ratings from the three SLPI rating team members were the same as or within one rating level after the second independent ratings for all scale levels except the following:

- Advanced Plus-Superior Plus range: Two (4.6%) of the 43 interviews. For these two interviews, raters for one team had second independent ratings of Intermediate Plus, Advanced, and Advanced Plus, and raters for the other team had second independent ratings of one Intermediate Plus and two Advanced Plus.

- Intermediate: One (3.2%) of the 31 interviews. This interview had second independent ratings of one Advanced and two Intermediate.

Table 1. Sign Language Proficiency Interview (SLPI; Newell, Caccamise, Boardman, & Holcomb, 1983) official ratings received by National Technical Institute for the Deaf SLPI interviewees from the AY 03–04 fall quarter through the AY 05–06 fall quarter (N = 157).

<table>
<thead>
<tr>
<th>SLPI official rating</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Plus-Superior Plus range</td>
<td>43</td>
<td>27.4</td>
</tr>
<tr>
<td>Advanced</td>
<td>15</td>
<td>9.6</td>
</tr>
<tr>
<td>Intermediate Plus</td>
<td>29</td>
<td>18.5</td>
</tr>
<tr>
<td>Intermediate</td>
<td>31</td>
<td>19.7</td>
</tr>
<tr>
<td>Survival Plus</td>
<td>21</td>
<td>13.4</td>
</tr>
<tr>
<td>Survival</td>
<td>14</td>
<td>8.9</td>
</tr>
<tr>
<td>Novice Plus</td>
<td>3</td>
<td>1.9</td>
</tr>
<tr>
<td>Novice</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>No functional skills</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>157</td>
<td>100.0</td>
</tr>
</tbody>
</table>

6The primary reasons that interviews may not be ratable are related to interviewer behavior. These reasons include the interviewee giving long, elaborated responses that are not clear but the interviewer does not request clarification and the interviewer using slow-to-moderate, simplified signing throughout an interview, resulting in raters not being able to evaluate an interviewee’s sign language comprehension skills.
Survival: Two (14.2%) of the 14 interviews. For these two interviews, one team had second independent ratings of Novice Plus, Survival, and Survival Plus, and one team had second independent ratings of two Novice Plus and one Survival Plus.

**Prenegotiation Rater Validity**

Table 4 compares the 471 first independent ratings by all raters (157 ratings × 3 raters per team = 471) and, when needed, their second independent ratings (21 ratings × 3 raters per team = 63) before analysis and negotiation, with the official rating arrived at after analysis and negotiation. The comparisons in this table summarize the results for all SLPI official ratings received by interviewees in this study. Of 471 first independent ratings, 2.5% were not within one level of official ratings; after second independent ratings, this percentage was reduced to 0.6%.

Table 5 provides a comparison of the raters’ first and second independent ratings with official ratings for each scale level. All ratings were the same as or within one rating level of official ratings except the following:

- **Advanced Plus-Superior Plus range:** Of 129 ratings, two raters (1.6%) on two different rating teams had second independent ratings of Intermediate Plus; the other two raters for these two teams had second independent ratings of Advanced Plus.

---

**Table 2.** For all SLPI rating levels/ranges, number and percentage of times that the first and second independent ratings of the 3 SLPI rating team members were (a) all the same, (b) all within one level of each other, and (c) all not within one level of each other.

<table>
<thead>
<tr>
<th>Independent ratings of SLPI raters</th>
<th>All 3 raters same ratings</th>
<th>All 3 raters within 1 level of each other</th>
<th>All 3 raters not within 1 level of each other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>First only</td>
<td>70</td>
<td>44.6</td>
<td>66</td>
<td>42.0</td>
</tr>
<tr>
<td>Second only</td>
<td>10</td>
<td>47.6</td>
<td>6</td>
<td>28.6</td>
</tr>
<tr>
<td>After second</td>
<td>80</td>
<td>50.9</td>
<td>72</td>
<td>45.9</td>
</tr>
</tbody>
</table>

**Note.** A second independent rating was done only if the first independent ratings of all 3 raters were not within one level of each other.

*This row combines the results for all of the second independent ratings plus all of the first independent ratings when second independent ratings were not done.

**Table 3.** For each SLPI rating level/range, number and percentage of times that the first and second independent ratings of the 3 SLPI rating team members were (a) all the same, (b) all within one level of each other, and (c) all not within one level of each other.

<table>
<thead>
<tr>
<th>SLPI official rating</th>
<th>First independent ratings of 3 rating team members</th>
<th>After second independent ratings of 3 rating team members</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within 1 rating level</td>
<td>Not within 1 rating level</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Advanced Plus-Superior Plus range</td>
<td>43</td>
<td>27.6</td>
</tr>
<tr>
<td>Advanced</td>
<td>15</td>
<td>9.6</td>
</tr>
<tr>
<td>Intermediate Plus</td>
<td>29</td>
<td>18.6</td>
</tr>
<tr>
<td>Intermediate</td>
<td>31</td>
<td>19.9</td>
</tr>
<tr>
<td>Survival Plus</td>
<td>21</td>
<td>13.4</td>
</tr>
<tr>
<td>Survival</td>
<td>14</td>
<td>9.0</td>
</tr>
<tr>
<td>Novice Plus</td>
<td>3</td>
<td>1.3</td>
</tr>
<tr>
<td>Novice</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>No functional skills</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>157</td>
<td>100</td>
</tr>
</tbody>
</table>

**Note.** A second independent rating was done only if the first independent ratings of all 3 raters were not within one level of each other.

*These columns combine the results for all of the second independent ratings plus all of the first independent ratings when second independent ratings were not done.

- **Survival:** Two (14.2%) of the 14 interviews. For these two interviews, one team had second independent ratings of Novice Plus, Survival, and Survival Plus, and one team had second independent ratings of two Novice Plus and one Survival Plus.
Table 4. For all SLPI rating levels, number and percentage of times that the first and second independent ratings of the 3 SLPI rater team members were (a) all the same as the official ratings, (b) all within one level of the official ratings, and (c) all not within one level of the official ratings.

<table>
<thead>
<tr>
<th>Independent ratings of SLPI raters</th>
<th>All 3 raters same as official ratings</th>
<th>All 3 raters within 1 level of official ratings</th>
<th>All 3 raters not within 1 level of official ratings</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>First only</td>
<td>347</td>
<td>73.7</td>
<td>112</td>
<td>23.8</td>
</tr>
<tr>
<td>Second only</td>
<td>45</td>
<td>71.4</td>
<td>15</td>
<td>23.8</td>
</tr>
<tr>
<td>After second</td>
<td>375</td>
<td>79.6</td>
<td>93</td>
<td>19.7</td>
</tr>
</tbody>
</table>

Note. A second independent rating was done only if the first independent ratings of all 3 raters were not within one level of each other.

Table 5. For Each SLPI rating level/range, number and percentage of times that the first and second independent ratings of the 3 SLPI rater team members were (a) all the same as the official ratings, (b) all within one level of the official ratings, and (c) all not within one level of the official ratings.

<table>
<thead>
<tr>
<th>SLPI official rating</th>
<th>Relationship of first independent ratings of raters to official ratings</th>
<th>Relationship after second independent ratings of raters to official ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Same</td>
<td>Within 1 rating level</td>
</tr>
<tr>
<td>Rating</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Advanced Plus-Superior Plus range</td>
<td>129</td>
<td>27.4</td>
</tr>
<tr>
<td>Advanced</td>
<td>45</td>
<td>9.6</td>
</tr>
<tr>
<td>Intermediate Plus</td>
<td>87</td>
<td>18.5</td>
</tr>
<tr>
<td>Intermediate</td>
<td>93</td>
<td>19.7</td>
</tr>
<tr>
<td>Survival Plus</td>
<td>63</td>
<td>13.4</td>
</tr>
<tr>
<td>Survival</td>
<td>42</td>
<td>8.9</td>
</tr>
<tr>
<td>Novice Plus</td>
<td>9</td>
<td>1.9</td>
</tr>
</tbody>
</table>

Note. A second independent rating was done only if the first independent ratings of all 3 raters were not within one level of each other.

- **Intermediate**: Of 93 ratings, one rater (1.1%) had a second independent rating of Advanced; the other two raters on this rating team had second independent ratings of Intermediate.

**DISCUSSION**

The SLPI 3-rater team procedure was designed to converge toward increasingly reliable and valid official ratings. This procedure involves independent ratings and negotiated discussion and review of interviewees’ interviews in relation to a standardized scaled linguistic performance profile that is grounded in a theory of language use and learning. The NTID reliability and validity data reported in this article provide clear evidence that the SLPI 3-rater team procedure successfully accomplishes that goal. Furthermore, the evidence presented here suggests that the SLPI 3-rater team procedure entails minimal influence of idiosyncratic construct-irrelevant variance on official ratings. In addition, the low number of interviews that resulted in no official rating suggests that the interview protocol used in this study produced valid data samples and provided outcomes that had clear validity in the self-reflective opinions of the interviewees.

**RESULTS AND CONCLUSION**

Reliability analyses of the SLPI 3-rater team procedure showed that for first independent ratings, more than 86% of the raters produced ratings that were within one rating level.
of those of the other members of their rating teams. When all three rating team members’ first independent ratings were not within one rating level of one another, adding a second independent rating before rater team discussion boosted the reliability of ratings within one rating level to more than 96%. These results confirm that the reliability for NTID SLPI raters’ independent ratings before negotiated discussion, review, and determination of interviewees’ official ratings was very high.

The results of the validity analysis for this study indicate relatively strong validity for raters’ prenegotiated independent ratings. Only 2.5% of the first independent ratings fell outside one level of the official ratings. Adding a second independent rating before rater team discussion when the first independent ratings of the three raters were not all within one level of one another reduced this percentage to 0.6%. These results suggest that the influence of idiosyncratic construct-irrelevant variance on construct validity before negotiation is minor, and the negotiation process is effective in reducing the impact of this variance on SLPI validity in the majority of rater-disagreement cases. The NTID reliability and validity data reported in this article, therefore, provide clear evidence that the SLPI 3-rater team procedure results in reliable and valid official ratings.

We caution, however, that the impact on rater performance of variations in interview conditions and of rater-specific sources of construct-irrelevant variance can vary widely across assessment programs and interviewee populations. Therefore, programs using the SLPI should conduct their own reliability and validity studies in order to confirm that idiosyncratic sources of construct-irrelevant variance do not significantly compromise the reliability and validity of their assessment results. As suggested in the introduction to this article, information obtained from such studies should be used to assist in planning and providing in-service training for SLPI team members.

ACKNOWLEDGMENTS

We gratefully acknowledge the following National Technical Institute for the Deaf (NTID) faculty and staff for their contributions to setting up and maintaining the database needed to conduct this study: Geoff Poor, Office of Communication Assessment Services Coordinator; Michele Steele, SLPI Staff Assistant; and Mitchell Berger, Sr. Systems Development & Support Services Analyst/Programmer.

We also acknowledge the motivation provided by Protase Woodford of the Educational Testing Services, Princeton, NJ, to the development of the SLPI. In 1980, at the Third National Symposium on Sign Language Research and Teaching, Woodford presented information about the Language/Oral Proficiency Interview (L/OPI), which is an interview technique that is used to assess spoken language skills (Woodford, 1982). As Woodford explained, the L/OPI was developed by the U.S. federal government after World War II in response to a need to have an assessment tool that could be used to assess current and potential government employees’ spoken communication skills for a variety of foreign languages. Based on Woodford’s presentation, several project teams began to explore the application of interview techniques to the assessment of sign language, including faculty from the College of Staten Island, Gallaudet University (then Gallaudet College), and NTID.

REFERENCES


Contact author: Frank C. Caccamise, Department of Research and Teacher Education, National Technical Institute for the Deaf, Rochester Institute of Technology, 96 Lomb Memorial Drive, Rochester, NY, 14623-5604. E-mail: fcaccamise@rit.edu.
## APPENDIX A. SIGN LANGUAGE PROFICIENCY INTERVIEW (SLPI) RATING SCALE

<table>
<thead>
<tr>
<th>RATING</th>
<th>CONSTRUCT DESCRIPTORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior Plus</td>
<td>Able to have a fully shared and natural conversation, with in-depth elaboration for both social and work topics. All aspects of signing are native-like.</td>
</tr>
<tr>
<td>Superior</td>
<td>Able to have a fully shared conversation, with in-depth elaboration for both social and work topics. Very broad sign language vocabulary, near native-like production and fluency, excellent use of sign language grammatical features, and excellent comprehension for normal signing rate.</td>
</tr>
<tr>
<td>Advanced Plus</td>
<td>Exhibits some superior level skills, but not all and not consistently.</td>
</tr>
<tr>
<td>Advanced</td>
<td>Able to have a generally shared conversation with good, spontaneous elaboration for both social and work topics. Broad sign language vocabulary knowledge and clear, accurate production of signs and fingerspelling at a normal/near-normal rate; occasional misproductions do not detract from conversational flow. Good use of many sign language grammatical features and comprehension good for normal signing rate.</td>
</tr>
<tr>
<td>Intermediate Plus</td>
<td>Exhibits some advanced level skills, but not all and not consistently.</td>
</tr>
<tr>
<td>Intermediate</td>
<td>Able to discuss with some confidence routine social and work topics within a conversational format with some elaboration; generally 3 to 5 sentences. Good knowledge and control of everyday/basic sign language vocabulary with some sign vocabulary errors. Fairly clear signing at a moderate signing rate with some sign misproductions. Fair use of some sign language grammatical features and fairly good comprehension for a moderate-to-normal signing rate; a few repetitions and rephrasing of questions may be needed.</td>
</tr>
<tr>
<td>Survival Plus</td>
<td>Exhibits some intermediate level skills, but not all and not consistently.</td>
</tr>
<tr>
<td>Survival</td>
<td>Able to discuss basic social and work topics with responses generally 1 to 3 sentences in length. Some knowledge of basic sign language vocabulary with many sign vocabulary and/or sign production errors. Slow-to-moderate signing rate. Basic use of a few sign language grammatical features. Fair comprehension for signing produced at a slow-to-moderate rate with some repetition and rephrasing.</td>
</tr>
<tr>
<td>Novice Plus</td>
<td>Exhibits some survival level skills, but not all and not consistently.</td>
</tr>
<tr>
<td>Novice</td>
<td>Able to provide single sign and some short phrase/sentence responses to basic questions signed at a slow-to-moderate rate with frequent repetition and rephrasing. Vocabulary primarily related to everyday work and/or social areas such as basic work-related signs, family members, basic objects, colors, numbers, names of weekdays, and time. Production and fluency characterized by many sign production errors and by a slow rate with frequent inappropriate pauses/ hesitations.</td>
</tr>
<tr>
<td>No Functional Skills</td>
<td>(May be) Able to provide short single sign and “primarily” fingerspelled responses to some basic questions signed at a slow rate with extensive repetition and rephrasing.</td>
</tr>
</tbody>
</table>

*Note.* Adapted from the U.S. Foreign Service Institute & ACTFL LPI Rating Scales by William Newell and Frank Caccamise, rev. ed., June, 2006. For all SLPI rating descriptors, the first statement (in bold type) is always a statement of American Sign Language (ASL) communicative functioning, with all remaining statements (regular type) being descriptors of ASL form (vocabulary, production, fluency, grammar, and comprehension).
APPENDIX B. SLPI INDIVIDUAL RATER WORKSHEET B

(USE WITH SLPI THREE AND TWO RATER TEAM PROCEDURES)

CANDIDATE: ______________________________________ RATING DATE: ____________________
INTERVIEWER: ____________________________________ RATER: __________________________

1. Functional Range: Above Intermediate or At Intermediate or Below Intermediate

2. Functional Descriptor:
   _____ Superior/Superior +: conversation shared and natural; in-depth elaboration
   _____ Advanced Plus
   _____ Advanced: generally fluent, shared, conversation; spontaneous elaboration
   _____ Intermediate Plus
   _____ Intermediate: generally responds in 3 to 5 sentences; conversation fairly shared for social and work topics
   _____ Survival Plus
   _____ Survival: generally 1 to 3 sentence responses with many questions by interviewer required
   _____ Novice Plus
   _____ Novice: one sign, short phrases, single sentence, memorized like signing

Both function and form considered for ratings #3 to #6.

3. Possible Final Ratings: ______________________________________________________________

4. Independent Rating after 1st viewing: _________________________________________________

5. Independent Rating after 2nd viewing; only done if 1st ratings not all within one level of each other: ______________________________________________________________

6. Independent Final Rating following discussion: __________________________________________

7. Linguistic Factors (Form):

   A. Vocabulary Knowledge
      How broad? ___________________________________________________________

      ________________ sign selection errors based on meaning:

      Overuse of ________________ initialized signs:

      ________________ fingerspelled words with standard signs:

   B. Production & C. Fluency
      How clear? ___________________ Rate ____________________

      Sign production errors (How many?; a few, some, etc.):

   D. Grammatical Features
      ________________ use of ________________ sign language grammatical features, including

      1) indexing, space, body shifts & eye gaze:

      2) sign verb movement directionality:

      3) facial expression & sign movement modification (size, degree, manner, & time):
4) repetition of sign verb movement:

5) repetition of sign noun movement and vertical & horizontal sweep:

6) number incorporation (may include use of timeline):

7) listing on non-dominant hand, FINISH, body shifts, & pauses:

8) sign word order (rh?; topic-comment; FINISH, MUST, CAN, & NOT at end of comments; conditional; Object-Subject-Verb):

9) classifiers:

10) non-manual signals (neg. & aff. head movements, yes-no?, wh-?, puffed cheeks, pursed lips/o-o, mm, pah, cs, th, clenched teeth):

11) other:

12) problems:

E. Comprehension

________________________________________________________________________________ comprehension of sign language produced at a
________________________________________________________________________________ rate, with ___________________repetition(s) and rephrasing(s).
## APPENDIX C. SLPI RATER DISCUSSION WORKSHEET B

(USE WITH SLPI THREE AND TWO RATER TEAM PROCEDURES)

| CANDIDATE: ____________________________________ | RATING DATE: ____________________ |
| INTERVIEWER: ____________________________________ | RATER: __________________________ |

1. Functional Range: Above Intermediate or At Intermediate or Below Intermediate

2. Functional Descriptor:

   - Superior/ Superior +: conversation shared and natural; in-depth elaboration
   - Advanced Plus
   - Advanced: generally fluent, shared, conversation; spontaneous elaboration
   - Intermediate Plus
   - Intermediate: generally responds in 3 to 5 sentences; conversation fairly shared for social and work topics
   - Survival Plus
   - Survival: generally 1 to 3 sentence responses with many questions by interviewer required
   - Novice Plus
   - Novice: one sign, short phrases, single sentence, memorized like signing
   - No Functional Skills: (may be) some one sign, single phrase responses

3. Independent Ratings:

   - A. After 1st Viewing __________________________
   - B. After 2nd Viewing __________________________
   - C. After Team Discussion __________________________

4. Linguistic Factors (Form): ________________________________________________________________

   **A. Vocabulary Knowledge**

   knowledge of basic? sign language vocabulary (EXAMPLES)

   with sign selection errors based on meaning (EXAMPLES)

   overuse of initialized signs (EXAMPLES)

   and fingerspelled words with standard signs (EXAMPLES)

   **B. Production & C. Fluency**

   production of signing at a rate, with sign production errors

   **D. Grammatical Features**

   use of sign language grammatical features, including

   1) indexing, space, body shifts & eye gaze:

   2) sign verb movement directionality:

   3) facial expression & sign movement modification (size, degree, manner, & time):

   4) repetition of sign verb movement:

   5) repetition of sign noun movement and vertical & horizontal sweep:
6) number incorporation (may include use of timeline):

7) listing on non-dominant hand, FINISH, body shifts, & pauses:

8) sign word order (rh?; topic-comment; FINISH, MUST, CAN, & NOT at end of comments;
   conditional; Object-Subject-Verb):

9) classifiers:

10) non-manual signals (neg. & aff. head movements, yes-no?, wh-?, puffed cheeks, pursed lips/o-o,
    mm, pah, cs, th, clenched teeth):

11) other:

12) problems:

**E. Comprehension**

________________________________________comprehension of sign language produced at a
________________________________________rate, with ________________repetition(s) and rephrasing(s).