Using Semantic Textual Similarity to Improve Back Translation and Validation of Instruments in American Sign Language (ASL)

Embree, J. A., Huber, J., Taylor, K.C., & Wilson, J. F.
Substance Abuse Resources & Disability Issues (SARDI) Program

Boonshoft School of Medicine
Wright State University

I. Introduction

Numerous barriers exist when attempting to provide culturally appropriate assessment instruments for persons who are Deaf. In addition to a scarcity of specialized instruments created in American Sign Language (ASL), validated versions of translated instruments are rare. This contributes to an increased likelihood of errors in screening, costs time and money for providers, and contributes to Deaf consumers increased disenfranchisement with pursuing treatment. The lack of culturally and linguistically accessible instruments is in large part due to the complexity and high cost of producing such instruments.

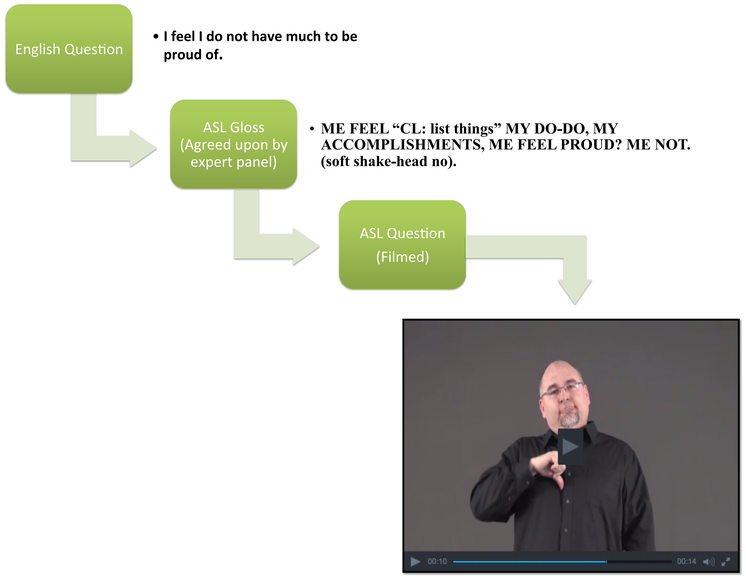
II. Population

Individuals (n = 30), both Deaf and Children of Deaf Adults (CODA), participated in back translation. Bi-lingual Deaf and CODAs were chosen because their native language is ASL and they have English fluency. Cognitive interviews were conducted with culturally Deaf (N = 20) people across the United States, and field-testing was done through a secure online portal (N = 450+).

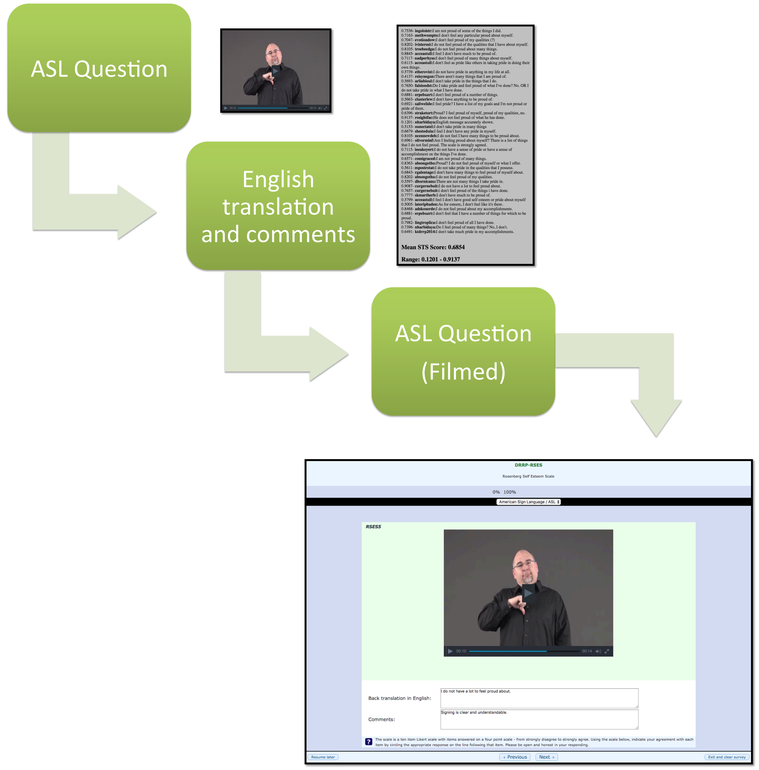
III. Procedure

English versions of instruments were translated, in a process called forward translation, into ASL gloss by a team of Culturally Deaf experts. This process was completed for the Satisfaction With Life Scale (SWLS), Rosenberg Self Esteem Scale (RSES), Beck Depression Inventory (BDI), Alcohol Use Disorders Identification Test (AUDIT), Drug Abuse Screening Test (DAST), Wide Range Interest and Occupation Test (WRIOT), Operational Network Evaluation Tool (O'NET), and Global Appraisal of Individual Needs (GAIN). Upon completion of the forward translation, another team of ASL experts, both Deaf and Hearing, recorded a video of each question in ASL. During this part of the process, ASL grammar and structure are fine-tuned. The back translation consists of two parts – to translate the ASL version of what is seen into English and to provide any comments with regards to ASL structure, grammar, and syntax. These translations are compared to the original text and the ASL gloss, using a semantic textual similarity (STS) analysis, combining distributional similarity and latent similarity analysis (LSA). Based on this analysis, the expert panel decides whether any of the questions need to be re-recorded, and the process is repeated. After all portions of an instrument are equivalent in meaning to the original, the videos continue to the process of cognitive interviewing and field-testing.

IV. Forward Translation



V. Back Translation



I feel I do not have much to be proud of.

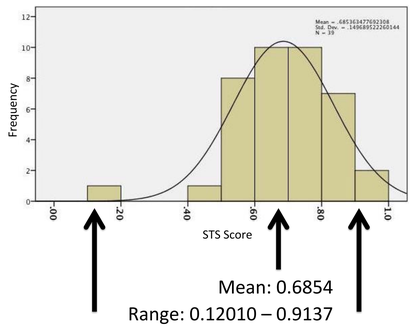
I feel I do not have much to be proud of.	I do not have a lot to feel proud about.	I don't take pride in many things.
1.0000	0.8835	0.5153

VI. Interpreting Scores

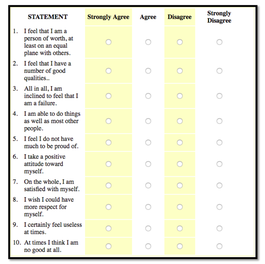
One of the reasons to use Latent Semantic Analysis (LSA) to measure semantic similarity is to improve analysis of back translation results and help decide which items need to be revised and reshot. Quantifying the results can reduce subjectivity and researcher bias, and aggregate scores can make it possible to look at large sets of data to understand how in item failed to perform so that it can be improved in future reshoots.

ASL Gloss → Back Translation
0.6238, *SD* = .1398

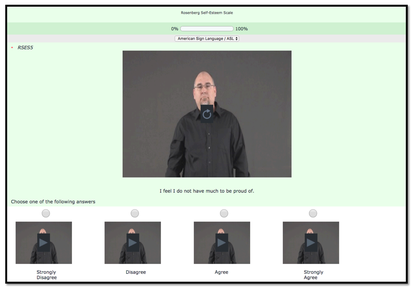
English → Back Translation
0.6854, *SD* = .1497



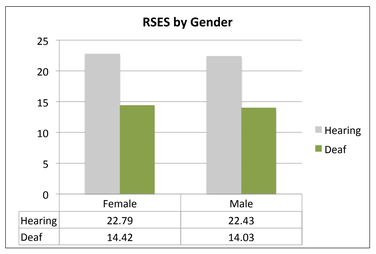
VII. Field Testing



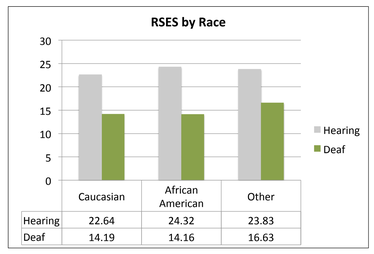
STATEMENT	Strongly Agree	Agree	Disagree	Strongly Disagree
1. I feel that I am a person of worth, at least on an equal plane with others.				
2. I feel that I have a number of good qualities..				
3. All in all, I am inclined to feel that I am a failure.				
4. I am able to do things as well as most other people.				
5. I feel I do not have much to be proud of.				
6. I take a positive attitude toward myself.				
7. On the whole, I am satisfied with myself.				
8. I wish I could have more respect for myself.				
9. I certainly feel useless at times.				
10. At times I think I am no good at all.				



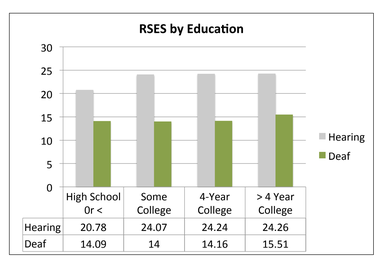
Field testing was conducted for five of the instruments thus far. The RSES was administered to Deaf participants (N = 466) in 45 states. Results were paired with scores from the RSES administered to the general US population (N = 503) in a study from 2010 (Sinclair, et al.), and representing 48 states. Overall scores for participants were in the range for low self esteem (14.23, SD = 5.975).



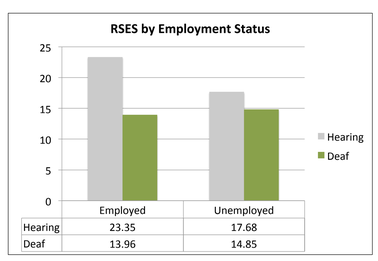
	Female	Male
Hearing	22.79	22.43
Deaf	14.42	14.03



	Caucasian	African American	Other
Hearing	22.64	24.32	23.83
Deaf	14.19	14.16	16.63



	High School Or <	Some College	4-Year College	> 4 Year College
Hearing	20.78	24.07	24.24	24.26
Deaf	14.09	14	14.16	15.51



	Employed	Unemployed
Hearing	23.35	17.68
Deaf	13.96	14.85

VIII. Results

Results suggest that STS scores can accurately predict similarity between ASL gloss and final back translations, as well as agreement among back-translators. STS scores comparing original English and final back translations are less promising, and suggest more complex conceptual differences between English and ASL.

IX. Conclusions

Although STS is not a substitution for a skilled panel of native-signers, the scores provided offer an objective way of comparing large samples of back translation content and reduce subjectivity among researchers during the process of choosing the most linguistically appropriate items. The possibility of crowd-sourcing back translation efforts and using natural language processing (NLP) holds particular promise in reducing costs and minimizing differences in regional signs that can contribute to reduced understanding.

X. References

Sinclair, S.J., Blais, M.A., Gansler, D.A., et al. (2010). Psychometric properties of the Rosenberg self-esteem scale: overall and across demographic groups living within the United States. *Evaluation & Health Professions, 33*, 56–80.

Funding for this program was made possible by a grant (90DP0067) from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR).