SIG4SIGN Webinar 2023
Developing rating scales for sign language assessments

Tobias Haug and Maria Tagarelli De Monte
SIG4SIGN Co-Convenors

The webinar was held online from 9.00 to 11:30 a.m. on November 22, 2023. Having the webinar online allowed people from all over Europe and Australia to participate as well. Among the attending people, there were participants from the U.S.A. and India as well. There were 203 people registered to the webinar, of which an average of 72 people have actually attended. At the end of the webinar, attendees and participants had some extra time to discuss about the topics being presented and, before leaving, they have been invited to respond to a brief questionnaire assessing their satisfaction with the webinar. We received 27 responses, all of which were happy with the webinar's contents and with the sign language interpreting service. Contents have been considered "very relevant" for participants' research or their professional practice (for a total of about 85% of respondents) and for "it covers very current topics" (55.56%). The webinar thus proves its value for the update and lifelong learning of researchers in the field of sign language testing and assessment as well as spoken language assessment and certifying agencies.

The webinar's length was extended from 1.5 to 2.5 hours, thanks to the funding received by the European Association for Language Testing and Assessment (EALTA) and by the University of Teacher Education in Special Needs, Institute for Language and Communication, Zurich (Tobias Haug's affiliation). This extended length was considered optimal by 85.19% of the respondents and while the online format remains the most interesting (preferred by 55.56% of the respondents), there is a growing number of people opting for hybrid or face-to-face meetings (25.93%, and 18.52%, respectively). Given the high number of preferences, the topic of next year's webinar will be "Creating CEFR-aligned tests/exams in sign languages".

The recording of the webinar, including sign language interpreting, was uploaded to the EALTA website.

Appendix: Program and abstracts of presentations

SIG FOR SIGNS WEBINAR SERIES
Developing rating scales for sign language assessments
November 22, 2023 from 09.00 to 11.30 CET
organized by
Tobias Haug and Maria Tagarelli De Monte

Program	
9.00 A.M.	Greetings and introduction
9.10 A.M.	Presentation 1 Considerations in rating scale development and validation by Ute Knoch, Melbourne University 20' speech, followed by 10' Q&A with the author
9.40 A.M.	Presentation 2 Development of rating scales for a Sentence Repetition Test for Swedish Sign Language by Krister Schönström, Stockholm University 20' speech, followed by 10' Q&A with the author
10.10 A.M.	Break
10.20 A.M.	Presentation 3 Development of a fluency rating scale for Swiss German Sign Language by Katja Tissi, Franz Holzknecht, & Tobias Haug, University of Teacher Education in Special Needs (HfH), Zurich; Alessia Battisti, University of Zurich; Nivja de Jong, Leiden University 20' speech, followed by 10' Q&A with the authors
10.50 A.M.	Final Q&A and discussion

The webinar will in English and International Signs, with the support of the interpreters of Overseas Interpreting.

The webinar will be recorded and then made available on the EALTA webpage and on the Facebook page of the EALTA - Special Interest Group for Signed Language Assessment (https://www.facebook.com/groups/1884613401752691/?mibextid=oMANbw)

Interpreting is funded by the European Association for Language Testing and Assessment (EALTA) and the University of Teacher Education in Special Needs, Institute for Language and Communication, Zurich.

Abstracts & Bios

Presentation 1
Considerations in rating scale development and validation
by Ute Knoch, Melbourne University

9.10 – 9.40 A.M.

Professor Ute Knoch is the Director of the Language Testing Research Centre at the University of Melbourne. Her research interests are in the areas of policy in language assessment, writing assessment, rating processes, assessing languages for academic and professional purposes, test validation and placement testing.

Find out more at https://findanexpert.unimelb.edu.au/profile/186949-ute-knoch

Abstract

Rating scale development in the field of language assessment is often considered in dichotomous ways: scale construction is assumed to be guided either by expert intuition or by drawing on performance data. Even though some have argued that rating scale development is rarely so easily classifiable, this dyadic view has dominated language testing research for over a decade.
In this talk, I will report on a study which aimed to identify the sources that typically impact rating scale design in published research. The study drew on a corpus of 42 published studies describing the design or revision of rating scales. The findings show that most rating scale development studies draw on a mixture of sources rather than either theory or performances. We identified 11 different sources of rating scale construct and present a model which divides these sources into test-internal and test-external sources. The paper argues that the source of the scale construct needs to align with the test purpose and score use from the outset, which was often not the case in the published research we examined. The most important contributions of the proposed model are that it offers an image of real-world rating scale development and gives rating scale developers a framework to consider prior to starting scale development or validation.

Presentation 2
Development of rating scales for a Sentence Repetition Test for Swedish Sign Language
by Krister Schönström, Stockholm University

9.40 - 10.10 A.M.

Krister Schönström is Associate professor at the Department of Linguistics at Stockholm University, Sweden. He received his PhD from Stockholm University in 2010 with a dissertation focusing on bilingualism in school-aged deaf children. His research interests include multilingualism of the deaf, language acquisition, and sign linguistics. He has been involved in several projects that involve sign language test development specifically STS-SRT and SignRepL2

Abstract

Within the field of sign language test development, sentence repetition tasks (SRTs) have gained popularity as a preferred test design for several sign languages. Particularly, ASL-SRT (Hauser et al., 2008) has served as the framework for some sign language tests, including BSL (Cormier et al., 2012), DGS (Kubus et al., 2015), DSGS (Haug et al., 2020), and Swedish Sign Language (STS) (Schönström & Hauser, 2022). Additionally, there exist sentence repetition task-based tests (not derived from ASL-SRT) for LIS (Rinaldi et al., 2018) and LSF (Bogliotti et al., 2020), among others. The primary rationale for this trend may be the sensitivity of SRTs in assessing overall language proficiency, encompassing sentence processing, reconstruction, and reproduction. Another factor contributing to this preference may be the practicality of these tests, characterized by their relatively brief administration and rating times.
I have been actively involved in the development of two distinct STS tests following the SRT paradigm: STS-SRT (Schönström & Hauser, 2022), which was crafted based on ASL-SRT, and SignRepL2 (Holmström et al., in press). These tests employ two different rating scales. STS-SRT utilizes a binary rating scale based on correct/incorrect responses, while SignRepL2 employs a more graduated rating scale comprising five points.
During my presentation, I will elaborate on these tests, specifically delving into the rating scales employed in SRTs, and engage in a discussion concerning the advantages and disadvantages of these rating scales.

Katja Tissi, Franz Holzknecht & Tobias Haug
University of Teacher Education in Special Needs (HfH),Zurich

Alessia Battisti
University of Zurich

Nivja de Jong
Leiden University

Presentation 3
Development of a fluency rating scale for Swiss German Sign Language
by Katja Tissi, Franz Holzknecht, & Tobias Haug, University of Teacher Education in Special Needs (HfH), Zurich; Alessia Battisti, University of Zurich; Nivja de Jong, Leiden University

10.20 – 10.50 A.M.

Abstract

In our presentation, we will report on the development and use of a fluency rating scale for Swiss German Sign Language (Deutschschweizerische Gebärdensprache, DSGS). The rating scale was developed by analyzing annotated productions of nine deaf L1 DSGS users, 10 L2 advanced DSGS users (interpreters), and 11 L2 DSGS beginning learners. The annotated data were based on a total of 162 performances elicited through tasks with and without preparation time. As more proficient users of a language are hypothesized to be more fluent than less proficient users, and as preparation time is hypothesized to lead to more fluent productions than absence of preparation time, the annotated productions of these performances can inform us on fluency aspects in sign language. Based on the statistical analysis of the annotated data, differences in terms of signing fluency between the three groups of participants were identified. These results informed the design of the following six rating criteria: (1) number of produced pauses, (2) length of pauses, (3) use of non-manuals during the production of pauses, (4) speed of signing, (5) repetitions, (6) and self-corrections. Each criterion is rated on a 6-point Likert scale. In a next step, three raters were first trained in using the newly developed rating scale before evaluating the productions of all study participants using the scale. A Many-Facet Rasch analysis showed that the three raters were able to apply the scale reliably, despite differences in rater severity. The model also reliably separated the six rating criteria, indicating that the individual criteria tap into different sub-constructs.