

Guidelines for sign language test development, evaluation, and use

Tobias Haug¹, Wolfgang Mann², Eveline Boers-Visker³, Jessica Contreras⁴, Charlotte Enns⁵,
Ros Herman⁶, & Katherine Rowley⁷

¹University of Applied Sciences for Special Needs Education, Switzerland

² University of Roehampton, UK

³Utrecht University of Applied Sciences & University of Amsterdam, The Netherlands

⁴University of Connecticut, USA

⁵University of Manitoba, Canada

⁶City, University of London, UK

⁷University College London, UK

Version 1.0

10.11.2016

Contact information:

tobias.haug@slas.ch

How to quote this document:

Haug, T., Mann, W., Boers-Visker, E., Contreras, J., Enns, C., Herman, R., & Rowley, K. (2016). *Guidelines for sign language test development, evaluation, and use*. Unpublished document, retrieved from <http://www.signlang-assessment.info/>.

Table of content

The guidelines are divided into the following sections:

- A) Introduction.....3
- B) Ethical issues.....4
- C) Issues regarding signed language test development.....5
- D) Testing deaf children in different contexts.....6
- E) Resources.....8
- F) Summaries of existing guidelines.....9

A) Introduction

Background of the guidelines

The following guidelines are the result of the Swiss National Science Foundation (SNSF) funded International Exploratory Workshop [Methodological Challenges and Technical Innovations in Signed Language Assessment](#) in Zurich, September 2015. These guidelines were generated from discussions between all workshop participants and compiled by the authors. Additional information is based on following, existing publications:

- Assessment Guidelines for Sign Language Interpreting Programmes
- [Different guidelines of the International Test Commission](#)
- [Sign Language Linguistics Society's \(SLLS\) Ethics Statement for Sign Language Research](#)

The workshop participants emphasized the importance to have these guidelines translated into International Sign. Due to budget constraints, it was not possible to provide such a translation, but this remains something we aim for in the future.

Who are these guidelines for?

The guidelines target Deaf and hearing researchers, test developers, test administrators, sign language teachers/instructors, and (if applicable) raters of sign language tests. When we use the term “sign language tests” we refer both to tests that assess adult learners of a sign language as a first or second/foreign language and/or tests that assess children with sign language as their first language. Even though most available sign language tests target children, the guidelines are equally applicable for professionals with an interest in sign language testing and assessment of adult learners of a sign language as a first or second/foreign language, for example, in sign language interpreter training programs.

How to use the guidelines?

The following guidelines address issues related specifically to sign language tests and testing of children since most of the existing guidelines focus on tests for adult learners. Links are provided to existing guidelines for test development, such as from the International Testing Commission (ITC), or the European Association of Language Testing and Assessment (EALTA), which include more general, construct-independent issues on (language) tests to provide additional/in-depth information. The guidelines stated here serve as a point of reference to develop, evaluate, and use tests, both for children or adult learners of a sign language. To investigate specific topics more in-depth, we recommend using existing guidelines (see *Additional resources and guidelines for (language) test development*) or refer to publications on sign language test development and adaptation (see *Selected references*).

B) Ethical issues in testing and working with Deaf communities

1. Ensure test administrators are qualified to use sign language tests, e.g., level of sign proficiency, adequate training in sign language assessment and interpretation of test results, etc.
2. Set and maintain high personal standards of competence in delivery of sign language assessments and interpretation of test results.
3. Keep up to date about relevant changes and advances relating to test use and development. These may include changes in legislation and policy, which may impact on tests and test use, and the importance of updating test norms.
4. Make sure that test materials (e.g., test instructions, items) and test data are kept securely at all times with limited access.
5. Respect copyright law and agreements that exist with respect to a test, including any prohibitions on the copying or transmission of materials in electronic or other forms to other people, whether qualified or otherwise.
6. Refrain from coaching individuals on actual test materials or other practice materials that might unfairly influence their test performance.
7. Treat test results confidentially and securely in the case of online delivered tests.
8. Make sure to obtain parental consent when testing children, and assent from children prior to testing. As part of the consent procedure, specify who will have access to results and define levels of confidentiality to individuals and their families before test administration.
9. Obtain any relevant consent prior to releasing test results to others.
10. Remove names and other personal identifiers from databases of results that are archived, for research use, development of norms or other statistical purposes.
11. Produce a reasoned justification for the use of tests.
12. Establish that the knowledge, skills, abilities, aptitudes or other characteristics, which the tests are intended to measure, are correlates of relevant behaviors in the context about which inferences are to be drawn.
13. Determine that the test's technical and user documentation provides sufficient information to enable evaluation of the following:
 - a. Coverage and representativeness of test content, representative norm groups, difficulty level of content, etc.
 - b. Accuracy of measurement and reliability demonstrated with respect to relevant populations.
 - c. Validity (demonstrated with respect to relevant populations) and relevance for the required use.
 - d. Freedom from systematic bias in relation to the intended test taker groups.
 - e. Practicality, including time required, costs, and resource needs.
 - f. Avoid judging a test solely on the basis of face value, test-user testimonials, or advice from those with a vested commercial interest.

C) Issues regarding sign language test development

In this section we address issues that are considered as important with particular regard to sign language tests.

1. Specific aspects for sign language test development

1. Involve native signers at each stage of the process (from development to dissemination). These should ideally include Deaf native signers with academic training or other relevant training or experience. This is motivated by the fact that for some sign languages, limited research is available. Therefore, input is required from native signers to ensure that test items appropriately reflect the structure of the particular sign language.
2. Involve Deaf/hearing people with different areas of expertise, such as linguists, psychologists, high-level interpreters, media design experts, computer programmers (for Web-/mobile-assisted sign language testing).
3. Make sure that all materials are age appropriate for the test-takers and are also visually clear and accessible to the target population, avoid high memory load, do not require written responses, etc.
4. Assure that the choice of elicitation techniques, item formats, test conventions, and procedures are readily accessible to all intended populations (e.g., children, adults).
5. Ascertain that item content and stimulus materials are familiar to all intended populations (e.g., images that are suitable for children might not be appropriate to be used with adults).
6. The language used in the directions and items themselves as well as in the handbook/manual should be appropriate for all cultural and language populations for whom the test or instrument is intended.
7. In case of a test adaptation, test developers/publishers should ensure that the adaptation process takes full account of linguistic and cultural differences among the populations for whom adapted versions of the test or instrument are intended. (See also *Selected references* for more information regarding sign language test adaptation.)

D) Testing Deaf children in different contexts

1. General issues testing Deaf children

1. When carrying out assessments, it is important to use different approaches to obtain as much information as possible about the child's language skills. This can be achieved using a range of methods:
 - a. Include both receptive and productive instruments to assess sign language.
 - b. Include different instruments whenever possible that focus on different aspects of sign languages, such as vocabulary, grammar, narrative skills.
 - c. Environment (carry out assessments in a range of settings, e.g., clinical, school and/or home settings). The child is likely to be most relaxed in a familiar setting and will most likely respond differently depending on the test environment.
 - d. It is important to realize that all testing will be "inauthentic", however we can try to achieve the closest to natural settings.
 - e. Try to use more than one sign language test, if available. Different tests may yield different results or test slightly different aspects of a child's language skills.
2. Appropriate standards for assessors:
 - a. High level of sign language skills, preferably fluent in the given sign language.
 - b. The ability to communicate flexibly to meet the needs of individual Deaf children (many Deaf children today have exposure to a sign language (e.g., BSL)/ sign systems (e.g., Signed English) skills) / sign accompanying spoken language (e.g., Sign Supported English).
 - c. Experience in working with Deaf children.
 - d. People involved in the assessment process (teamwork): We stress the importance of Deaf and hearing teamwork when carrying out language assessments as each member of the team will have a different set of skills.
3. Aims of assessment:
 - a. The aims of the assessment need to be clear.
 - b. It is important that the assessment is actually testing what you want to test, e.g., vocabulary knowledge vs reading skills (validity).
4. Interaction:
 - a. It is vital to include live interaction with the child you are testing. Computer-based assessments are becoming more and more popular due to its efficiency but it does not always capture the child's ability to use language interactively.
5. General data protection:
 - a. Ensure that the data collected is secure and not identifiable.
 - b. Describe what happens to the data when the project ends or if the person responsible for the data leaves.
 - c. Follow data protection laws of your country and keep up-to-date with changes in law in your country.
 - d. Ensure that appropriate ethical approval is obtained before collecting and storing personal data.
 - e. Ensure that individuals and/or their parent/guardian consents to collected data about them being stored and to any specific use being made of their personal data including test scores.
6. Responsibility:

- a. The people who carry out sign language assessments are usually the first people to get a good understanding of the child being assessed. Assessors have a responsibility to ensure the outcomes of assessments are comprehensible and can be shared with the people who work with the child on a daily basis.

2.1 Clinical application

1. Follow recommendations outlined in general guidelines.
2. Collection of data:
 - a. Create a profile of the child from several data sources.
 - b. Use different data sources (previous reports about the child, video recordings of the child, results of previous assessments).
3. Use checklists/questionnaires to collect relevant background information about the child, e.g., preferred language, language experience/exposure, education, family information, medical history, etc.
4. Determine the “age appropriate” language level of the child: This can only be measured reliably when age norms are available for the respective sign language. Due to heterogeneity regarding the age of exposure, quality, and frequency of linguistic input that deaf children receive, it can be hard to determine what “age appropriate” means. Deaf children in hearing families comprise the majority of the deaf population and often acquire a sign language later in life. However, this should not be taken as the benchmark for sign language acquisition, which should be based on children in deaf families.
5. Targets/recommendations: Based on the information obtained from the assessment, a set of targets and recommendations should be identified for the child if the child is underperforming for their age.

2.3 Education (schools)

1. Learning outcomes: There should be several learning outcomes from assessing the child.
2. Schools may aim to assess both *academic achievement* and *language proficiency* (the two aren't necessarily the same). Assessing academic achievement covers curriculum-related aspects, which may not be appropriate for all deaf children. Assessing language proficiency tests level of language fluency and it should be possible to see individual progress over time.
3. Portfolio or Language Passport:
 - a. It may be useful for schools to create a language portfolio for each child, which will include reports, observations, scores from assessments collected over time, and including language background.
 - b. Marking grid, it may be useful to create a scoring grid to assess language competency and fluency that can be used cross-linguistically (i.e. for sign and spoken languages?) and this should link up to learning outcomes.
4. Involvement of children in the assessments.
 - a. Older children can be involved in understanding the results of their assessments.
 - b. They can be involved in discussions about new targets for them to achieve and how they can achieve those new targets.
 - c. It would be excellent for them to see their own progress over time and reflect on their own language skills.

E) Resources

Sign language tests resources

ProSign: Sign Languages for Professional Purposes (ECML project, 2012-15): [Resource site on the CEFR and sign languages, with some information on assessment](#)

Overview of existing signed language tests: www.signlang-assessment.info

Final report of the SNSF International Exploratory Workshop on “[Methodological challenges and technical innovations in signed language assessment](#)”

Additional resources and guidelines for (language) test development

European Association for Language Testing and Assessment (EALTA): [EALTA Guidelines for Good Practice in Language Testing and Assessment](#)

These guidelines are available in **35** different languages, which can be downloaded here: <http://www.ealta.eu.org/guidelines.htm>

Association of Language Testing and Assessment (ALTE): [Manual for Language Test Development and Examination. For use with the CEFR](#)

International Language Testing Association (ILTA): [ILTA Guidelines for Practice](#)

International Testing Commission (ITC): [Different guidelines on test development \(not specifically referring to language tests\)](#)

Selected references on sign language test development and adaptation and the assessment process

Haug, T. (2011). [Approaching sign language test construction: Adaptation of the German Sign language receptive skills test](#). *Journal of Deaf Studies and Deaf Education*, 16(3), 343–361. <http://doi.org/10.1093/deafed/enq062>

Haug, T., & Mann, W. (2007). [Adapting tests of sign language assessment for other sign languages - A review of linguistic, cultural, and psychometric problems](#). *Journal of Deaf Studies and Deaf Education*, 13(1), 138–147. <http://doi.org/10.1093/deafed/enm027>

Herman, R. (2015). Language assessment in deaf learners. In H. Knoors & M. Marschark (Eds.), *Educating deaf learners: Creating a global evidence base* (pp. 197-212). New York, NY: Oxford University Press.

Mann, W., & Haug, T. (2014). Mapping out guidelines for the development and use of sign language assessment: Some critical issues, comments and suggestions. In D. Quinto-Pozos (Ed.), *Multilingual aspects of signed language communication and disorders* (pp. 123–139). Bristol, UK: Multilingual Matters.

References (used for this guidelines)

Leeson, L., Calle Alberdi, L., & Bown, S. (2013). [Assessment guidelines for sign language interpreting training programmes](#). European Forum of Sign Language Interpreters.

F) Summary of existing guidelines

1. Issues on test use (based on International Test Commission (2013), [ITC Guidelines for Test Use](#))

1.1 General and specific guidelines

Knowledge, understanding, and skill

Knowledge, understanding and skill underpin all test administrator competencies. The nature of their content and level of detail are likely to vary between countries, areas of application and as a function of the level of competence required to use a test. The Guidelines do not contain detailed descriptions of these. However, when applying the Guidelines for use in specific situations the relevant knowledge, skills, abilities and other personal characteristics will need to be specified. This specification is part of the process of contextualization, through which generic guidelines are developed into specific standards. The main areas descriptions of knowledge, understanding and skills need to cover include the following.

Relevant declarative knowledge, including:

1. Knowledge of basic psychometric principles and procedures, and the technical requirements of tests (e.g., reliability, validity, standardization).
2. Knowledge of tests and measurement sufficient to inform the choice of tests and the interpretation of test results relevant to one's area of practice.

Instrumental knowledge and skills, including:

1. Knowledge and skills relating to specific assessment procedures or instruments, including the use of computer-based assessment procedures.
2. Specialized knowledge of and practitioner skills associated with using those tests that are within one's repertoire of assessment tools.
3. Knowledge and understanding of the construct or constructs underlying test scores, where this is important if valid inferences are to be drawn from the test results.

General personal task-related skills, including:

1. The performance of relevant activities such as test administration, reporting, and the provision of feedback to test takers and other stakeholders (e.g., parents when the test taker is a child).
2. Oral and written communication skills sufficient for the proper preparation of test takers, test administration, the reporting of test results, and for interaction with relevant others (e.g., parents, or organizational policy makers).
3. Interpersonal skills sufficient for the proper preparation of test takers, the administration of tests, and the provision of feedback of test results.

Task management skills, including:

1. Knowledge of codes of conduct and good practice relating to the use of tests, test data, the provision of feedback, the production and storage of reports, the storage of and responsibility for test materials and test data.
2. Knowledge of the social, cultural, and political context in which the test is being used, and the ways in which such factors might affect the results, their interpretation and the use to which they are put.

1.2 Issues in testing and analysis

The competent test administrator will make all reasonable efforts to make necessary

preparations for the testing session which include:

1. Providing relevant parties in a timely manner with clear information concerning the purpose of testing and the procedures to be followed.
2. Advising test takers of the linguistic or dialectic groups for which the test is considered appropriate.
3. Explaining clearly to test takers their rights and responsibilities.
4. Obtaining consent of test takers or their legal guardians or representatives prior to testing.
5. Explaining in cases where testing is optional, the consequences of taking or not taking the test to relevant parties so that they can make an informed choice.

The competent test administrator will make the necessary practical arrangements by ensuring that:

1. Locations and facilities for testing have been arranged well in advance, and the physical environment is accessible, safe, quiet, free from distractions and appropriate for the purpose.
2. Staff who will be involved in the administration are familiar with administering the test and have the language skills required.
3. Appropriate arrangements have been made for the testing of people with additional needs, e.g., visually accessible stimuli, slower presentation rate of test items, support in place to assist with test takers who have additional needs.

Competent test administrators will administer the tests properly by:

1. Establishing rapport by welcoming test-takers and briefing them in a positive fashion.
2. Acting to reduce test-taker anxiety and avoid creating or reinforcing unnecessary anxiety, e.g. by not referring to a 'test' with a young child.
3. Ensuring test-takers have the materials they require for taking the test before it begins administering tests under appropriate supervised conditions.
4. Wherever possible, administering test instructions in the primary language of the test takers, even where the test content is designed to provide evidence of knowledge or skills in a non-primary language.
5. Providing adequate time for examples to be completed and offering additional practice opportunities if needed.
6. Ensuring that anyone assisting the administration has had proper training.
7. Following the standardized procedures for test administration.

Competent test administrators will score and **analyze** test results accurately by:

1. Following carefully the standardized procedures for scoring.
2. Carrying out appropriate raw score conversions to other relevant types of scaled score.
3. Ensuring that invalid conclusions are not drawn from comparisons of scores with norms that are not relevant to the people being tested.

Competent test administrators will interpret results appropriately by:

1. Having a good professional understanding of the test's theoretical or conceptual basis, and guidance on the use and interpretation of the scale scores.
2. Having a good understanding of the scales used, the characteristics of the norm or comparison groups, and the limitations of the scores.
3. Using appropriate norm or comparison groups where available and interpreting results in the light of available information about the person being tested (including age, gender, schooling, culture, language experience/exposure, etc.) with due regard for the technical limitations of the test, and the assessment context.
4. Avoiding over-generalizing the results of one test to traits or human characteristics that

are not measured by the test.

5. Considering each scale's reliability, error of measurement and other qualities, which may have artificially lowered or raised results when interpreting scores.
6. Giving due consideration to the available evidence of validity, with respect to the construct being measured for members of the test takers' relevant demographic groups (e.g., cultural, age, social class, and gender groups.)
7. Taking into account and reporting any individual or group variations from standard procedures in test administration.

Competent test administrators will communicate the results clearly and accurately by:

1. Ensuring that the technical and linguistic levels of any reports are appropriate for the level of understanding of the recipients, including the test-taker, where appropriate.
2. Making clear when the test data represent just one source of information and should therefore always be considered in conjunction with other information.
3. Explaining how the importance of the test results should be weighted in relation to other information about the person being assessed.
4. Using a form and structure for a report that is appropriate to the context of the assessment

2. Technical issues (partly based on International Test Commission (2010), [A Test-Taker's Guide to Technology-Based Testing](#))

Competent test administrators will:

1. Ensure the system (e.g., computers to be used for test administration) meet the hardware and software requirements defined by the test developer prior to administering any tests. For web-based tests, this includes making sure there are no browser (type of browser like Firefox or Chrome and version of browser) or operation system (e.g., Windows, Linux, Apple) restrictions.
2. Make sure to have high-speed Internet access when the test includes sign language videos.
3. In the case where the test includes online video-recording, check if a webcam is available and works with the used browser. Make sure to have appropriate lighting conditions and background when sign language videos are recorded.
4. Know how to access technical support (online as well as at the testing site) in case of any problems.
5. Have read the instructions and looked at any supporting materials needed to successfully complete the test. This includes completing any online tutorials (where available).
6. Be familiar with operating the computer or other device and the software functions required by the test.
7. Inform test takers about the purpose of the test, the content included in the test, and the type of test (e.g., picture-matching, multiple choice), where appropriate.
8. Provide clear instructions how to take the test.
9. Make sure the test is administered in a comfortable and quiet environment with minimal distraction and appropriate lighting conditions.
10. Make sure to follow the standard administration procedures.
11. Make sure that the test takers do not have access to other programs on the computer to look up information while being tested (i.e. they should only have access to the test).
12. Make sure that the test takers cannot store/copy the test on devices they bring (e.g., external hard-drive.)
13. If possible, make sure that the computer has some sort of logbook file to prevent a test takers claim that „the recording is lost“.

14. Make sure that test takers do not see each other as information in sign language is easily passed through over a distance, for example when assessing adult learner in a sign language interpreter training program.

In cases where the test can be self-administered, test takers should expect:

1. Technical support contact information in case of any problems.
2. Clear instructions on how to take the test, access the test website, log on/log off to the test.
3. Proper maintenance of data security, in particular if data is transferred over the Internet.
4. Proper maintenance of data confidentiality through use of encryption or passwords. Access to any test data should be restricted to those who are qualified and authorized.
5. Use standard encryption technology for online testing (SSL, as know from online banking) and check with any legal issues of a test provider.